BIOLOGICAL REVIEWS

of the

Cambridge Philosophical Society

Edited
by
H. MUNRO FOX

VOLUME 19

CAMBRIDGE
AT THE UNIVERSITY PRESS
1944

LONDON
CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, N.W. 1

BOMBAY, CALCUTTA, MADRAS: MACMILLAN

PRINTED IN GREAT BRITAIN

CONTENTS

No. 1, January 1944

No. 2, April 1944

No. 3, July 1944

No. 4, October 1944

INDEX OF AUTHORS

PHOTOSYNTHESIS, PHOTOREDUCTION AND DARK REDUCTION OF CARBON DIOXIDE IN CERTAIN ALGAE

By HANS GAFFRON, From the Department of Chemistry, Fels Fund, University of Chicago, Chicago, Illinois

(*Received* 23 *May* 1943)

I. PROBLEM AND DEFINITIONS

Among the green algae a number of species belonging to the genera *Scenedesmus*, *Rhaphidium* and *Ankistrodesmus* change their metabolism into a type different from that of chlorophyllous plants and previously found only in some bacteria as the result of anaerobic treatment with hydrogen (G., 1939*a*, 1940*a*; 1942*a*). Normally these algae, like all green plants, reduce carbon dioxide with the simultaneous evolution of an equal volume of oxygen, but after the hydrogen treatment or 'adaptation' they perform the same task with the simultaneous absorption of two volumes of hydrogen. The reduction with hydrogen can proceed in two ways. Either photochemically with the aid of the light energy absorbed by chlorophyll or in the dark with the aid of the energy derived from the oxidation of hydrogen to water (Gaffron, 1942*b*). The fact that photosynthesis, photoreduction and dark reduction of carbon dioxide can occur in the same cell is important for the understanding of the mechanism of photosynthesis.

Enough observations have accumulated at present to make it clear that only a rather complicated sequence of reactions, comparable to those revealed in the studies on respiration, will be sufficient to represent the true course of photosynthesis in plants. The separation of the process of photosynthesis into its partial reactions has been initiated in two ways. (1) Preparations of chloroplasts, which have lost the capacity to reduce carbon dioxide, have been shown to produce free oxygen upon illumination in solutions of organic iron salts (Hill, 1939; Hill & Scarisbrick, 1940). This proves that oxygen can be produced photochemically aside from any reduction of carbon dioxide. (2) The photochemical reduction of carbon dioxide with molecular hydrogen in green algae proves that carbon dioxide reduction can be reversibly separated from oxygen evolution. These observations obviously indicate the existence of two partially overlapping reaction cycles. This article is a review of the experimental results on the reactions with hydrogen: they confirm the general theoretical aspects of the problem of photosynthesis as presented in two recent reviews (Franck & Gaffron, 1941; van Niel, 1941) and allow us to draw a diagram showing the interrelation of the different types of carbon dioxide reduction and the most probable sequence of experimentally recognizable partial reactions. The terms used most frequently in this article are defined as follows: *Photosynthesis:* the over-all process of photochemical carbon dioxide reduction coupled with the production of free molecular oxygen. *Photochemical reaction:* the absorption of light energy by chlorophyll and the chemical reactions so closely connected with it that at present they cannot be observed separately. *Photoreduction:* the over-all process of photochemical carbon dioxide reduction coupled with the utilization of hydrogen, so that water is formed instead of oxygen. *Dark reactions:* all non-photochemical reactions essential to photosynthesis which are not a part of the photochemical process as defined above. Despite the obvious fact that the reactions involved in photosynthesis form a continuous cycle as far as the catalytic system is concerned, it is practical to distinguish the reactions which precede the photochemical process from those which follow it. *Fixation of carbon dioxide:* one of the dark reactions, namely, the reversible binding of carbon dioxide in the chloroplast preceding the photochemical reduction. *Dark reduction* (*or chemo-synthesis*): the non-photochemical reduction of carbon dioxide promoted by an energy-yielding metabolic reaction.

II. ADAPTATION OF *SCENEDESMUS* AND SIMILAR ALGAE TO A METABOLISM INCLUDING MOLECULAR HYDROGEN

(1) *Plant material and method.* Metabolic reactions with molecular hydrogen have been observed so far in several species of unicellular green algae belonging to the genera *Scenedesmus*, *Ankistrodesmus* and *Rhaphidium*. Of these some species of *Scenedesmus* are easiest to cultivate in quantity and nearly as indifferent to changes in the surrounding medium as *Chlorella pyrenoidosa* which may be called the 'standard organism' for studies of photosynthesis. Equal ability for an anaerobic metabolism with hydrogen does not necessarily imply other physio-

logical similarities in two species of *Scenedesmus*, which may differ greatly, for example, in the sensitivity of their respiratory metabolism toward cyanide (Gaffron, 1939*b*). No extensive search has been made for other plants capable of reactions with hydrogen, and it is likely that more will be found. The available strains of algae were kept and used as absolutely pure cultures. The results reported have been gained with Warburg's manometric method for determining the metabolic gas exchange in living cells. Any other method generally used to measure photosynthesis may be equally suitable. A method which would allow one to trace simultaneously the exchange of each of the gases involved, oxygen, hydrogen and carbon dioxide, would be clearly superior.

(2) *Appearance of a 'hydrogenase'.* Hydrogen has been used very often as a means to replace air in experiments with plant material when anaerobic conditions were desired. No effects of this gas on the metabolism of plants other than those observed in an atmosphere of nitrogen have been reported. Hydrogen has been considered, therefore, as an inert gas for the plant.

It is common knowledge that green plants begin to ferment as soon as the partial pressure of oxygen in the medium drops below that necessary to maintain the 'Pasteur effect' (inhibition of fermentation by oxygen). In the dark a suspension of *Scenedesmus* in water or in an acid buffer begins to produce carbon dioxide when placed in an atmosphere of nitrogen. No other gas is liberated provided the anaerobic period lasts not longer than an hour at room temperature. The same is true if nitrogen is replaced by hydrogen. A protracted anaerobiosis, however, changes the picture. After about 2 hr. at room temperature in hydrogen *Scenedesmus* and similar algae start to *absorb* hydrogen. When kept during the same time in nitrogen these organisms begin to *develop* hydrogen. This can be demonstrated by the reduction of methylene blue on a platinum surface or by absorption by palladium black. A complete but short removal of oxygen is not sufficient to induce the metabolic change. The adaptation period, however, is shortened if one uses unphysiologically high temperatures. At 36° C. the absorption or evolution of hydrogen by *Scenedesmus* begins as soon as all traces of oxygen have been removed. At this high temperature, however, the photochemical abilities of the algae are easily damaged or destroyed, whereas at 20° C. the algae survive unchanged several days of an anaerobiosis in the dark.

The development of the hydrogen metabolism in our algae is somewhat similar to the appearance of formic hydrogenlyase in *Bacillus coli* after transfer into a medium containing formate (Stephenson, 1939). The long incubation period can be explained in two ways. Either a new enzyme system has to be synthesized, or an inactive enzyme becomes active by reduction. The first assumption fits the observations on the metabolism of *B. coli* but not those on *Scenedesmus*. If we attribute the appearance of an active hydrogenase system to a reduction process, we imply that it has a high negative oxido-reduction potential. This agrees with the observations on the return to aerobic conditions described in § V. It seems that the hydrogen-transferring system is the first to be re-oxidized at the return to aerobic conditions.

How the presence of hydrogen during a long anaerobic incubation period influences the activities of the hydrogenase system is shown in the following figures: 100 mg. of cells (wet weight) exchanged in about 12 hr. (*a*) in nitrogen: $+130$ mm.^{3} CO_2, $+96$ mm.^{3} non-volatile acids, $+37$ mm.^{3} H_2; (*b*) in hydrogen: $+127$ mm.^{3} CO_2, $+96$ mm.^{3} non-volatile acids, -207 mm.^{3} H_2. The rate of the ordinary fermentation, producing carbon dioxide and non-volatile acids, is not affected by the partial pressure of hydrogen. The hydrogen exchange, however, depends on it. Since hydrogen can be absorbed in the dark in considerable quantities acceptors for it must either be present at the beginning of the anaerobic period or be produced during fermentation. Hence part of the hydrogen formed during fermentation in nitrogen may be taken up by such acceptors. To date no simple stoichiometric relationship between the hydrogen and the other products of fermentation has been obtained in *Scenedesmus* (Gaffron & Rubin, 1942; Rubin, 1941).

In order to observe a measurable production of hydrogen it is not necessary to add a substrate. The source of hydrogen is an unknown reserve substance in the cell. As can be expected the rate of the hydrogen evolution declines with time. Starved cells, which have been allowed to respire a long time in the dark, show a low rate of fermentation, while a high rate is found with cells which previously have assimilated carbon dioxide in the light for several hours. Of numerous organic substances added to the cell suspension only glucose was found to enhance immediately all phases of the anaerobic metabolism. Yet the increased production of hydrogen (in an atmosphere of nitrogen) remains rather small as compared with the quantity of hydrogen available in form of hydrogen donors. One conspicuous result of adding glucose is a rise in the percentage of lactic acid among the products of fermentation.

Summing up we may say that during the dark anaerobic adaptation period a hydrogenase system comes into play capable of liberating as well as absorbing molecular hydrogen. Which way the hydrogen is transferred depends on the partial pressure of this gas and the relative quantities and oxido-reduction potentials of hydrogen donors within the cell. The capacity to exchange hydrogen, acquired by anaerobic incubation, affects (as we shall see below) the course of other metabolic reactions in the cell, especially that of photosynthesis.

(3) *The effect of poisons upon the adaptation reac-*

tions. If small concentrations of *cyanide* ($10^{-4}M$) are present in the algal suspension, no adaptation occurs. Replacing the air above the cell suspension by nitrogen or hydrogen now only makes the algae ferment in the ordinary way. In *Scenedesmus* there are, thus, three reactions, at least, sensitive to cyanide: respiration, photosynthesis and the 'adaptation'. The last is the most sensitive. Complete failure to attain the 'reduced state' of the hydrogenase system can often be observed with cyanide concentrations which scarcely inhibit any other metabolic process. This effect of cyanide indicates that a heavy metal catalyst is involved in the adaptation reaction and possibly in the activity of the hydrogenase. We assume that the oxidized form of the hydrogenase is inactive and capable of associating itself with cyanide. Hence no adaptation occurs as long as small amounts of oxygen are present or when cyanide prevents the oxidized form from being reduced.

Hydroxylamine is a poison of particular interest because at low concentrations ($10^{-4}M$) it does not interfere with respiration, yet affects photosynthesis more strongly even than cyanide. The adaptation reaction was found to be sensitive to hydroxylamine, although not more so than photosynthesis. Practically the same results as with hydroxylamine have been obtained with 1, 10-*phenanthroline*. Phenanthroline forms heavy metal complex salts, particularly with divalent iron. Otherwise its properties are absolutely different from those of hydroxylamine.

One point needs to be emphasized. We can say that a poison inhibits specifically only the adaptation reaction if its effect depends upon whether it has been added before or after adaptation. This holds for cyanide and hydroxylamine or 1, 10-phenanthroline, but not for two other specific poisons widely used in cell physiology: dinitrophenol and carbon monoxide. In our algae, *dinitrophenol* is a poison which inhibits respiration and photosynthesis as well as the reactions of the hydrogenase system. The effects found upon adding dinitrophenol before or after adaptation do not differ sufficiently to indicate clearly a specific inhibition of the adaptation reaction. *Carbon monoxide* has only a slight influence upon the respiration of these algae and none at all upon photosynthesis. All activities of the hydrogenase system under anaerobic conditions are, however, strongly affected. Special experiments need still to be made in order to prove a specific inhibition of the adaptation reaction.

III. PHOTOREDUCTION

(1) *Photochemical reduction of carbon dioxide with molecular hydrogen.* When cells of *Scenedesmus* (or *Rhaphidium*, or *Ankistrodesmus*, etc.) are subjected to an anaerobic incubation as described, not only their fermentative metabolism is changed, but also the course of their photosynthetic reactions. Upon illumination in the presence of carbon dioxide and hydrogen these algae do not photosynthesize normally, i.e. with the evolution of oxygen, but absorb hydrogen together with carbon dioxide. The process is now described not by equation (*a*) but by equation (*b*):

$$(a)\quad n\,[CO_2+4H_2O+Xh\nu] \rightarrow (CH_2O)_n+n\,[3H_2O+O_2],$$

$$(b)\quad n\,[CO_2+2H_2+Xh\nu] \rightarrow (CH_2O)_n+n.H_2O,$$

where $(CH_2O)_n$ means an unknown carbohydrate. This reaction has been observed before only in some purple bacteria (Roelofsen, 1935; Gaffron, 1935*a*, 1935*b*) which also reduce carbon dioxide photochemically with various organic and inorganic hydrogen donors under strictly anaerobic conditions. These bacteria are unable to liberate molecular oxygen.

Equation (*b*), however, can be verified quantitatively only under stationary conditions. The transition from the stationary state in the dark to that in the light requires a certain time, an 'induction period'.

(2) *Induction phenomena.* It is well known that when plants are irradiated suddenly with intense light after a dark period the full rate of photosynthesis is reached only after an 'induction period' lasting from one to several minutes (Aufdemgarten, 1939; McAlister & Myers, 1939). The induction loss, that is, the amount of carbon dioxide which would have been reduced if no initial inhibition existed, is proportional to the final rate of photosynthesis. It becomes negligible at low light intensities; the length of the period is more or less independent of the intensity. Photoreduction with hydrogen also does not begin immediately at its full rate. In this case, however, the length of the induction period decreases if the intensity of the radiation increases, and the induction loss is more or less the same at all light intensities, at least if we compare algae from the same culture. (Grown differently, or suspended in different media, the algae exhibit quite different induction losses.) The anaerobic induction, therefore, can be explained by the presence of a certain amount of cellular material which has to be dehydrogenated photochemically before a steady rate of photoreduction with hydrogen is attained (Gaffron, 1940*a*). This assumption agrees well with the occasional photochemical liberation of hydrogen (§ IV) and with the fact that other hydrogen donors, like glucose, added to the medium may diminish the rate of absorption of molecular hydrogen (§ III (6)).

A phenomenon complementary to the induction period occurs when the light is turned off: the absorption of hydrogen does not stop immediately but continues for a few minutes at a rapidly decreasing rate. If hydrogen donors are present in the cell which at the beginning of a light period supply hydrogen to the photochemical system, it is not

surprising to find that afterwards in the dark these substances act as hydrogen acceptors. Actually the absorption of hydrogen in the dark continues as long as the cells are metabolizing (§ I).

(3) *The stationary state: influence of the light intensity.* In presence of an excess of hydrogen and carbon dioxide photoreduction proceeds at a steady rate once the induction period is over. This is true, however, only for low light intensities which under aerobic conditions would scarcely produce more than a compensation of respiration. Stronger radiation destroys the capacity of the plants to absorb hydrogen and causes them to revert to normal photosynthesis. This interesting reversion is described in detail in § V.

Normally, the rate of photosynthesis rises proportionally to the light intensity up to fairly high values of the latter (2000–10,000 foot candles) before it becomes 'light saturated', that is, before the curve representing rate versus intensity bends parallel to the abscissa. Because of the early reversion of photoreduction to photosynthesis it is not possible to obtain such a *light saturation* for the reaction with hydrogen. It is generally assumed that light saturation occurs when an enzymatic reaction connected with the photochemical processes attains its maximum rate (Franck & Herzfeld, 1941). In photoreduction the rate of hydrogen transfer by the hydrogenase system apparently limits the over-all rate, because the reversion occurs at light intensities which are so much lower than that necessary to obtain the aerobic light saturation (cf. § VII). The conditions governing the rate of photoreduction resemble those existing with photosynthesis in presence of inhibitory concentrations of cyanide. This can be shown by experiments in which the plants are illuminated not continuously but by a sequence of very short but extremely bright light flashes.

(4) *Flash saturation.* It has been known for many years (Warburg, 1928) that cyanide inhibits one of the enzymatic dark reactions of photosynthesis. Cyanide decreases the maximum rate obtainable at saturation intensities, but scarcely interferes with the rate of photosynthesis at low light intensities where the rate is still proportional to the intensity of radiation. It was quite natural to assume that the enzymatic reaction responsible for limiting the rate at light saturation is the one which is affected by cyanide. The experiments with flashing light, however, have shown recently that cyanide inhibits a dark reaction that is not normally limiting at light saturation (Weller & Franck, 1941; Rieke & Gaffron, 1943). In presence of cyanide this dark reaction becomes so slow that it sets the pace for the over-all process. It is now generally believed that this cyanide-sensitive reaction is the thermal fixation of carbon dioxide. Emerson & Arnold (1932) demonstrated that flashing light is the proper means to separate what we call the photochemical reaction from some of the slower enzymatic reactions. Each flash must be so short (10^{-4} sec.) that the photosynthetic process cannot run through more than one catalytic cycle while the flash lasts. The flashes must be spaced so far apart (10^{-2}–10^{-1} sec.) that even the slowest partial reaction initiated by the absorption of light ends before the next flash appears.

If the so-called photochemical reaction consisted only of the absorption of light by chlorophyll, the yield of photosynthesis calculated for a single flash should rise with the intensity of the flash to very high values. Actually Emerson & Arnold found that the yield per flash attains soon a maximum and stops increasing further with the intensity of the flash: a *flash saturation* has been reached. This result proves that our definition of the photochemical reaction includes at least one, if not more, chemical or enzymatic reactions which are much slower than the absorption of a light quantum.

If we illuminate algae at regular intervals with flashes producing flash saturation, the momentary intensity is higher than necessary to produce light saturation in continuous illumination. Yet the total amount of photosynthetic products formed during a given time of flashing illumination can be very small because it depends on the average intensity. The average intensity is high with short intervals between flashes, low with long ones. The shortest interval between flashes which still allows a maximum yield per flash is normally determined by the time required for the reactions comprising the photochemical process. Those dark reactions which either furnish the photosensitive intermediates or deal with the products of the photochemical reaction are fast enough not to interfere with the flash saturation. If we inhibit, however, one of these dark reactions with cyanide, we create a bottleneck allowing only a certain amount of material to be produced in a certain period of time. Under these conditions the maximum yield per flash appears to be affected by the poison so long as the rate of the over-all reaction is determined by the rate of the inhibited reaction and not by the average light intensity. By spacing out the flashes we decrease the over-all rate of photosynthesis, the effect of the bottleneck vanishes and the yield per flash increases to the normal flash saturation value (Emerson & Arnold, 1932). We conclude that the photochemical process itself is not affected by cyanide. It can be shown that this conclusion is correct by repeating the flashing light experiments with cyanide in a manner which allows one to keep the average intensity constant despite great changes in the intervals between flashes. The trick consists in spacing a given number of flashes per second unevenly in groups of two or four, with short dark intervals within each group and long ones between the groups (Rieke & Gaffron, 1943).

The same experiment can be done with photoreduction. Light of excessive intensity does not cause a reversion to normal photosynthesis, provided the radiation is applied in short flashes alternating

with comparatively long dark periods. Here again the total amount of light absorbed and, consequently, the amount of photochemical products formed during the unit of time is small. Under these circumstances the maximum yield of photoreduction *per flash* is the same as under aerobic conditions. What matters for the undisturbed continuation of photoreduction is the total amount of primary photochemical products formed in the total elapsed time, but not the true intensity of the radiation by which they have been formed. If the amount of photoproducts per unit time is so small that they can be supplied by the flow of carbon dioxide despite cyanide inhibition, or by the flow of hydrogen despite the restricted capacity of the hydrogenase system, neither an inhibition of photosynthesis by cyanide nor a reversion from photoreduction is found even with the most intense flashes.

(5) *Quantum yield and assimilatory quotient* (Dutton & Manning, 1941; Emerson & Lewis, 1939, 1942; Manning, Stauffer, Duggar & Daniels, 1938). Since the flash saturation is the same for photoreduction and photosynthesis it is very probable that the adaptation leaves the photochemical mechanism unchanged. Hence it is not too surprising to find that the quantum yield of the reduction of carbon dioxide in *Scenedesmus* is also the same whether measured during photosynthesis or during photoreduction. The yield is 0·1. Approximately 10 quanta of visible light are necessary to reduce one molecule of carbon dioxide with either the liberation of one molecule of oxygen or the absorption of two molecules of hydrogen (Rieke, unpublished).

The knowledge of the ratio $+\Delta O_2/-\Delta CO_2$ (assimilatory quotient) during the period of observation has been of paramount importance for the calculation of the quantum yield (Emerson & Lewis, 1939, 1942). The above-mentioned quantum yield of 0·1 has been calculated on the assumption, confirmed by experimental determinations (Gaffron, 1940*a*), that the ratio $-\Delta H_2/-\Delta CO_2$ is equal to 2. Such determinations encounter one difficulty: while quantitative measurements of photosynthesis are complicated by the presence of respiration, quantitative measurements of photoreduction have to be corrected for the continuous absorption of hydrogen caused by a non-photochemical fermentation as described in § II (2). After such a correction had been applied twelve determinations gave values of the assimilatory quotient ranging from 1·84 to 2·17, with an average of 1·97. This shows that photoreduction can be represented by equation (*b*) in § III (1).

(6) *The competition of other hydrogen donors.* In § II we have seen that, in the dark, the adapted algae absorb or deliver hydrogen according to conditions influencing the hydrogenase system. Since in the light this system transfers hydrogen to the photochemical mechanism it is evident that hydrogen originating from intracellular material could be used for the photochemical reduction of carbon dioxide. It is a question of relative reaction velocities whether the internal hydrogen donors participate in the reduction of carbon dioxide in competition with molecular hydrogen. The hydrogen liberation during fermentation declines rapidly in cells containing little or no reserve material, but is sustained for many hours in cells fed with glucose. Correspondingly, we find that the presence of glucose in the suspension medium depresses the uptake of hydrogen by the photoreducing algae. The quantum yield, as measured by the rate of hydrogen absorption, drops 20–50%. Even a complete inhibition of hydrogen uptake in the light has occasionally been observed, although not with glucose but with yeast autolysate. Since these effects can be observed in bicarbonate solutions it is likely that they are not caused by an accumulation of free organic acids in toxic concentrations (Noack, Pirson & Michels, 1939). The duration of anaerobic incubation in presence of organic material determines to a certain extent the 'inhibition' effect which this material can exert upon the absorption of hydrogen in the light. This is readily understood since not glucose but some metabolic derivative is the actual hydrogen donor. Once oxygen has been readmitted to the assimilating algae the 'inhibition' by the organic material vanishes. The cells continue with photosynthesis in the same manner as the 'uninhibited' controls.

The rate of the hydrogen fermentation in nitrogen in the dark is very much slower than the rate of photoreduction with molecular hydrogen when reversion sets in. Thus, in order to produce the effects described above, the competition of the internally supplied with the gaseous hydrogen evidently cannot occur after the former has been released as free gas. We have to assume that the two converging lines of hydrogen transfer compete at a place in the enzymatic chain where the slow conversion from bound to free hydrogen does not yet enter into the picture.

(7) *Effect of specific poisons on photoreduction* (Gaffron, 1942*a*). The analysis of complex metabolic processes by the use of specific poisons, a method which has proved very fruitful in the studies of respiration and fermentation, met with difficulties soon after it was introduced by Warburg into the study of photosynthesis twenty years ago. The reason was that under the influence of poisons photosynthesis becomes inhibited in its entirety: there was no accumulation of intermediates nor any deviation from the normal chemistry. The separation of oxygen liberation from carbon dioxide reduction now provides new opportunities for an analysis of the mechanism of photosynthesis with the aid of specific poisons. In these experiments it is important to add the poison after complete adaptation to the anaerobic metabolism, as pointed out in § II (3), because the adaptation reaction itself is very sensitive to *cyanide* or to *hydroxylamine*. Photoreduction is far less sensitive.

There is no apparent distinction between the inhibitory effects of cyanide and of hydroxylamine on normal photosynthesis, whereas a decisive difference becomes apparent in photoreduction. Hydroxylamine, which even in small concentrations is a powerful inhibitor for photosynthesis, has, in the same concentrations, no observable effect upon photoreduction. Cyanide, on the other hand, inhibits equally both reactions. Thus hydroxylamine (in small concentrations) must be a specific poison for the oxygen-liberating system, in contrast to cyanide which apparently interferes with the reduction of carbon dioxide as well as with the activity of the hydrogenase.

In as complicated a process as photosynthesis a poison may often inhibit more than one partial reaction. Obviously the effect of the poison will be traced first to the most sensitive reaction. Whether the less sensitive ones will be identified too depends mainly on being able to let the process continue after the first inhibited partial reaction has been eliminated. The effect of hydroxylamine on photosynthesis has been attributed above to the oxygen-liberating system. Yet photoreduction is only relatively insensitive to hydroxylamine. Larger concentrations ($M/800$ to $M/30$) diminish the rate of hydrogen absorption to about one-half or less in the course of a few hours. The inhibition never becomes complete.

In contrast to *cyanide* and *hydroxylamine*, which are assumed to react exclusively with heavy metal catalysts, the *dinitrophenols* are supposed to combine with, and inhibit the activity of, some protein taking part in the transfer of hydrogen (Haas, Harrer & Hogness, 1942). Experiments comparing the effect of 2, 4-*dinitrophenol* upon the aerobic and anaerobic reduction of carbon dioxide revealed that both reactions are equally susceptible to this poison. In all probability, therefore, it is the transfer of hydrogen to carbon dioxide which is inhibited. The action of dinitrophenols on the algal metabolism shows the traits already described for other organisms (Krahl & Clowes, 1938, 1940).

The only poison which was observed to interfere quite specifically with photoreduction (but not with photosynthesis) is *carbon monoxide*. In an atmosphere of carbon monoxide and hydrogen the algae revert to normal photosynthesis at a light intensity at which photoreduction in a nitrogen-hydrogen mixture continues for hours. Carbon monoxide appears to have a special affinity to the hydrogenase. I may cite a few other examples. The complicated photochemical metabolism of some sulphur-purple-bacteria (Thiorhodaceae) containing a hydrogenase is changed by the presence of carbon monoxide (Gaffron, 1935*b*) so that less carbon dioxide is assimilated. The production of hydrogen during the fermentation of butyric acid bacteria is inhibited by this gas (Kempner, 1933; Kempner & Kubowitz, 1933). In *Azotobacter* (Wilson & Wilson, 1943) carbon monoxide inhibits the reaction between oxygen and hydrogen much more than respiration.

(8) *Absence of growth during photoreduction.* Purple bacteria grow abundantly under anaerobic conditions, but all attempts to cultivate the adapted algae anaerobically have been unsuccessful. Of two samples of *Scenedesmus* suspended in the same nutrient medium, irradiated with the same intensity, and initially assimilating carbon dioxide at the same rate, only the one in air showed formation of chlorophyll, multiplication and hence an increase in the rate of carbon dioxide absorption; the sample adapted to photoreduction showed after a few hours a decline in the rate of photochemical activity which continued for several days of uninterrupted illumination until the rate was only a tenth of its initial value. The cells of this sample exhibited no signs of deterioration, yet they could not be 'revived' quickly by transfer to aerobic conditions. The ability to reduce carbon dioxide remains, however, practically undiminished for many days, despite anaerobic conditions, if the cells are kept in the dark. It is not surprising that the rate of respiration immediately after an anaerobic period is much greater than normal. Products of photoreduction and fermentation may accumulate during the anaerobic period which are readily oxidized once oxygen has again been admitted.

IV. THE PHOTOCHEMICAL EVOLUTION OF HYDROGEN

(1) *Photochemical activity in absence of carbon dioxide.* The algae of the *Scenedesmus* type adapt themselves to a metabolism with hydrogen in an atmosphere of nitrogen as readily as in an atmosphere of hydrogen (see § II), the difference being only that in nitrogen the dark fermentation is accompanied by a *liberation* of small amounts of hydrogen. Upon illumination the algae adapted in nitrogen change to photoreduction, but only for a short period of time, since the partial pressure of hydrogen developed during the adaptation period is very low. After this the algae revert to photosynthesis. Course and duration of these transitions depend on the light intensity and on the conditions of the plants. A large supply of internal hydrogen donors tends to stabilize the 'reduced state'.

The anaerobic conditions, however, remain perfectly stable and the hydrogenase system active despite the absence of hydrogen if the other substrate, carbon dioxide, is also absent. This can be accomplished by providing a solution of alkali which absorbs all carbon dioxide from the gas phase above the cell suspension. The irradiated algae can make use of only that small part of the carbon dioxide formed in fermentation which does not escape from the suspension. In the light, therefore, there will be neither an appreciable photoreduction, nor a return to photosynthesis, nor photoxidation. Hence the

light energy absorbed by the cells kept in pure nitrogen could be expected to be transformed into heat without producing any measurable metabolic reaction. Actually, however, the irradiated algae develop a new photochemical reaction. They produce free hydrogen at a rate which initially is at least ten times the rate of the dark fermentation. The gas production continues for hours, though at a rapidly declining rate. The nature of the gas is easily established by its reaction with palladium black or methylene blue in contact with platinum. It does not react with an alkaline solution of pyrogallol, hence does not contain oxygen. Finally, if the nitrogen is replaced by hydrogen containing a few per cent of carbon dioxide, the algae turn immediately to photoreduction without any period of adaptation.

The photochemical evolution of hydrogen depends upon the presence of suitable hydrogen donors in the cell. As in the case of the dark fermentation the yield of this reaction is improved by adding glucose to the medium, or by a preceding period of intense photosynthesis. Theoretically important is the fact that the light effect is restricted to an increase of the production of hydrogen alone. The rates at which carbon dioxide and fixed acids are produced remain the same as in the dark control. It follows from this that the effect of light in this peculiar reaction cannot be described as a photochemical acceleration of the dark fermentation.

(2) *Separation of the photochemical evolution of hydrogen from the dark fermentation.* Evidence that the evolution of hydrogen in light is a reaction different from the release of hydrogen in the dark is found in the action of poisons. With *dinitrophenol* it is possible to inhibit largely or even completely the appearance of hydrogen during the fermentation in the dark. If the poisoned algae are now irradiated, the photochemical production of hydrogen takes place at the normal or even an accelerated rate. Dinitrophenol, however, does not inhibit all types of fermentation reactions. With this poison there is a change from the hydrogen fermentation to one producing mainly lactic acid. This is also known in *Bacillus butyricus* (Kempner, 1933; Kempner & Kubowitz, 1933).

In the poisoned algae the rate of the photochemical evolution of hydrogen usually exceeds that of the controls. To understand this apparent stimulation we need only remember that dinitrophenol inhibits photosynthesis and photoreduction. The cell fermenting in nitrogen has some carbon dioxide at its disposal even if carbon dioxide is continually removed from the gas phase by absorption with potassium hydroxide. These small amounts of carbon dioxide react with part of the hydrogen which becomes available for reduction by irradiation. Dinitrophenol inhibits the reduction, thereby allowing more hydrogen to escape. Thus a photochemical production of hydrogen can be demonstrated even without removing carbon dioxide provided the cell suspension contains enough dinitrophenol. The absolute rate of the photochemical hydrogen production is very small compared with the saturation rate of normal photosynthesis in continuous light. It is smaller even than the rate of photoreduction with hydrogen, which itself is at best only about three times the rate of respiration.

V. THE REVERSION FROM PHOTOREDUCTION TO PHOTOSYNTHESIS

(1) *Intermediates causing reversion.* It has been said above (§ III (3)) that no true light saturation of photoreduction in continuous light can be obtained because the algae return to normal photosynthesis if the intensity of the radiation surpasses a certain value (which varies with internal factors of the particular strain of algae). The time required for this 'reversion' is a function of the 'excess' light intensity. A part of the reversion process must be the reversal of the adaptation reaction described in § II, since once the normal evolution of oxygen has started it is not possible to restore photoreduction merely by reducing the intensity of illumination. To achieve this one has to repeat the original adaptation procedure. The first indication that the algae are beginning to revert to normal photosynthesis is the slowing down of the hydrogen uptake. If the light is turned off during this period the reversion may be completed in the dark, or the plant may recover its ability for photoreduction in a comparatively short time. In the latter case more hydrogen than usual is absorbed in the dark after an exposure.

The ability for photoreduction can be destroyed by light even if the plants are prevented by poisoning from liberating any molecular oxygen. This indicates that the reversion is due to an accumulation of an *intermediate* and not to the production of free oxygen. It is plausible that a precursor to molecular oxygen, if not reduced in time, may oxidize and inactivate the hydrogenase system. Aerobically the same intermediate would be decomposed by the oxygen-liberating system. We may call the hypothetical intermediary photoproduct with oxidizing properties 'photoperoxide' or simply 'peroxide'. In discussing the influence of specific poisons upon the reversion reaction it will be shown how necessary it is to assume the existence not only of this but also of at least one other intermediate between water and the eventual release of molecular oxygen (cf. § VII (1)).

(2) *Reversion in presence of hydroxylamine.* In small concentrations hydroxylamine is a strong poison for photosynthesis and hardly an inhibitor for photoreduction (cf. § III (7)). Consequently in the presence of hydroxylamine the photoperoxides should accumulate under the influence of excessive light exactly as in unpoisoned cells until they oxidize the hydrogenase system. This actually happens if the concentration of hydroxylamine is small. With a greater hydroxylamine concentration, however, the

reversion reaction is inhibited or retarded. In other words, the mechanism of photoreduction can be protected against reversion to photosynthesis by the presence of larger amounts of hydroxylamine. Algae adapted to photoreduction and poisoned with hydroxylamine continue to absorb hydrogen for hours, even if the intensity of irradiation is such that it would otherwise bring about the reversion to photosynthesis in a few minutes. The 'protection' caused by hydroxylamine, however, is not absolute. The higher the light intensity, the more hydroxylamine must be present to prevent the reversion. Once reversion has occurred, despite the poison, the algae stop all measurable photochemical activity because the 'protective' concentrations of the poison are more than sufficient to inhibit aerobic photosynthesis completely.

(3) *Light saturation in algae poisoned with hydroxylamine.* We have seen in § III (3) that it is not possible to obtain a light saturation curve of photoreduction in continuous light because of the reversion to photosynthesis. If we prevent the reversion by sufficient hydroxylamine, the problem of light saturation can be taken up again. With increasing intensity the rate of photoreduction rises indeed above that at the normal reversion threshold, but it never attains the high rate of photosynthesis at saturation. At most the increase is 100%, even at the intensities which would saturate photosynthesis. This low saturation rate can be explained in at least two ways. (1) Hydroxylamine not only prevents the reversion reaction but it also inhibits at this high concentration the stationary rate of photoreduction (§ III (7)). The saturation then could be an effect of the poison. (2) The transfer of hydrogen by the hydrogenase is in itself a slow process, and the saturation value may be a measure of the maximum capacity of this system. Special experiments showed that the rate of photoreduction is clearly a function of the hydroxylamine concentrations at low intensities, whereas at light saturation different hydroxylamine concentrations hardly have an influence upon the maximum rate. We believe, therefore, that the saturation is determined by the capacity of the hydrogenase system (cf. § III (3)).

(4) *Reversion in presence of o-phenanthroline.* Hydroxylamine is a substance which readily reacts with keto and aldehyde groups in organic compounds, thereby forming oximes. The concentrations of hydroxylamine used in some of the experiments described, particularly in those where photoreduction was protected against a reversion, have been so high that oxime formation with intermediates rather than the specific inhibition of a catalyst might have been the cause of the effects observed. To decide which type of reaction occurs when the cell is poisoned a search has been made among the substances capable of forming complexes with heavy metals or with heavy metal catalysts. It was found that 1, 10-*phenanthroline*, a substance quite different from hydroxylamine and rather indifferent to carbonyl groups, could replace hydroxylamine as a poison in all the effects described: inhibition of photosynthesis, of adaptation, of photoreduction and of the reversion reaction. The difference in the effective concentrations characteristic of the inhibition with hydroxylamine was also found in the experiments with phenanthroline. Oxime formation, therefore, should be excluded as an explanation of the effects of hydroxylamine in photosynthesis.

(5) *Reversion in presence of cyanide.* Reasons have been given (§ III (4, 7)) why the effect of cyanide on photosynthesis is supposed to be an inhibition of the non-photochemical carbon dioxide fixation. Cyanide, therefore, should diminish the rate of photoreduction to the same extent as that of photosynthesis. This is indeed the case, but we should expect in addition that larger concentrations of cyanide have a protective effect upon the 'reduced' state and prevent a reversion to photosynthesis by decreasing the stationary concentration of the oxidized intermediate photoproducts. The experiments, however, show the opposite. Presence of cyanide produces a reversion to photosynthesis at lower intensities and faster than observed with the unpoisoned controls. In *Scenedesmus* the accelerated reversion can be very marked with concentrations of cyanide that do not interfere much with the rate of photosynthesis at the same limiting light intensity. The action of cyanide is truly antagonistic to that of hydroxylamine. Hydroxylamine protects the normal as well as the cyanide-poisoned algae against a quick reversion. In the latter case, however, the protection lasts only for a restricted period; slowly but steadily the rate of photoreduction declines until the cyanide effect has won.

(6) *Reversion in presence of dinitrophenol.* The effect of dinitrophenol is worth mentioning as presenting an interesting contrast to the action of cyanide and of hydroxylamine (or 1, 10-phenanthroline). This poison offers some protection against the reversion by excessive light; but it is obtained only at the price of a very strong inhibition of photoreduction. In other words, dinitrophenol regulates the flow of the oxidized photoproducts at the source. Its effect is thus similar to that of lack of carbon dioxide or the lowering of the light intensity. Increasing the rate of photoreduction by a more intense irradiation in the algae poisoned with dinitrophenol leads to a reversion when the rate approaches that which causes a reversion in the unpoisoned algae. After reversion in presence of dinitrophenol photosynthesis proceeds at an inhibited rate as determined by the concentration of the poison.

VI. DARK REDUCTION OF CARBON DIOXIDE COUPLED WITH THE OXYHYDROGEN REACTION

(1) *Influence of the partial pressure of oxygen.* Traces of oxygen are sufficient to prevent an adaptation to hydrogen, but measurable amounts of oxygen are tolerated *after* complete adaptation. This tolerance is due to the fact that the oxygen is now rapidly reduced to water. Many bacteria are known to catalyse the 'Knallgas', or oxyhydrogen reaction (Stephenson, 1939; Lipmann, 1943; Wilson & Wilson, 1943). In some of these organisms the oxyhydrogen reaction is coupled with a reduction of carbon dioxide (Ruhland, 1924). In *Scenedesmus* (and the other physiologically similar green algae), too, the oxyhydrogen reaction can induce a reduction of carbon dioxide in the dark (Gaffron, 1940*c*, 1942*b*).

The velocity at which adapted algae start to absorb oxygen together with hydrogen is proportional to the partial pressure of the oxygen. If, however, this pressure rises above a certain threshold, the algae revert to aerobic conditions; all reactions with hydrogen disappear and normal respiration sets in. This happens at partial pressures between 1 and 2% of an atmosphere, but the critical pressure varies greatly with internal conditions of the organisms. The maximum rate of the stable oxyhydrogen reaction corresponds closely to the maximum rate of photoreduction just before reversion. Photoreduction can be superimposed on the oxyhydrogen reaction, provided the total rate of hydrogen absorption does not surpass the maximum rate of each separate reaction (Franck & G., 1941, Table III). This constitutes a further support of the contention that the reversion reaction is always caused by an insufficient rate of hydrogen transfer. For quantitative experiments it is necessary to adjust the initial partial pressure of oxygen so low that no reversion occurs. Otherwise it makes little difference whether, *ceteris paribus*, the partial pressure of oxygen above the suspension is 0·3, 0·6 or 1·0% of an atmosphere. The total amount of gas absorbed is proportional to the oxygen partial pressure. The curves obtained from two different experiments representing the amount of oxygen absorbed as a function of time can be superimposed accurately simply by multiplying the ordinates with a factor given by the ratio of the partial pressures of oxygen used in these experiments.

(2) *The influence of carbon dioxide.* The rate at which the algae absorb hydrogen after the addition of a certain quantity of oxygen as well as the total amount of gas absorbed is determined by the presence or absence of carbon dioxide. If carbon dioxide is present, some of it disappears together with twice the volume of hydrogen, in addition to the amount of hydrogen used up in the formation of water. In absence of carbon dioxide not enough hydrogen is absorbed to convert the disappearing oxygen into water. In a large number of experiments a stoichiometric utilization of hydrogen and oxygen was observed only rarely. Numerous measurements of the oxyhydrogen reaction have shown that in presence of carbon dioxide the process corresponds in most cases to the equation

$$2O_2 + CO_2 + 6H_2 \rightarrow \underset{\text{carbohydrate}}{(CH_2O)} + 5H_2O, \quad (c)$$

whereas in complete absence of carbon dioxide the reaction fits the equation

$$O_2 + H_2 \rightarrow \begin{matrix}\text{—OH}\\ \text{—OH}\end{matrix} \text{ (an unknown 'peroxide').} \quad (d)$$

In other words, a complete reduction of oxygen to water

$$O_2 + 2H_2 \rightarrow 2H_2O \quad (e)$$

seems to occur only in the presence of carbon dioxide, either coupled with the reduction of CO_2, as in (*a*), or alone if the reduction of carbon dioxide is inhibited by specific poisons (see §§ VI (3); VII (3)). We have to emphasize that the presence of carbon dioxide enables the reduction of oxygen to proceed to completion, even if carbon dioxide itself is prevented from being reduced.

The important conclusion to be drawn from these experiments is that carbon dioxide plays a double part in the mechanism of the oxyhydrogen reaction. It is not only the substrate in a coupled oxido-reduction which proceeds with a remarkable efficiency (0·5 mol. CO_2 per mol. of O_2), but its presence or absence decides which course the very reduction of oxygen will take. If carbon dioxide is absent, only one molecule of hydrogen is absorbed (as a rule) per molecule of oxygen. The observable reaction proceeds only to an intermediate state, the 'peroxide-level', as indicated in equation (*d*).

(3) *The effect of specific poisons on the oxyhydrogen reaction.* The experiments on photoreduction reveal that cyanide and hydroxylamine have opposite effects (§§ III, V). The smallest amounts of cyanide favour the return to aerobic conditions, whereas hydroxylamine protects the adapted state against the return to photosynthesis. To determine whether the dark reduction of carbon dioxide and the photochemical process show any resemblance the influence of cyanide and of hydroxylamine upon the oxyhydrogen reaction has been studied. It was found that in presence of oxygen, in quantities which are harmless to the reduced state in the unpoisoned cells, a concentration of *cyanide* which inhibits neither respiration nor photosynthesis to any important degree causes an inactivation of the hydrogenase system. The cyanide effect is different in presence or absence of carbon dioxide. Reaction (*c*) is inhibited immediately in contrast to reaction (*d*) which continues up to an hour despite the presence of small amounts of cyanide ($10^{-4}M$). The final inactivation of the hydrogenase system by oxygen appears to be inevitable once cyanide is present. Only the time

interval required varies with the cyanide concentration. With much cyanide the oxyhydrogen reaction stops almost at once.

Hydroxylamine, on the other hand, has no effect at all upon the oxyhydrogen reaction in concentrations which inhibit normal photosynthesis very strongly ($1–2 \times 10^{-4} M$). Only if the concentration of this poison is increased so much that it clearly diminishes the rate of photoreduction in the same sample of algae (that is from $10^{-3} M$ upward) may one find a partial inhibition: reaction (*c*) changes into reaction (*e*). The coupled reduction of carbon dioxide disappears under the influence of hydroxylamine, while the formation of water continues. Cases have been observed, however, where $10^{-3} M$ hydroxylamine had no effect at all and the oxyhydrogen reaction continued undisturbed according to equation (*c*). This shows that unknown internal factors play an important part. There is no doubt that hydroxylamine always penetrates into the chloroplast, because upon illumination one invariably finds all the specific effects of the poison on photoreduction or on photosynthesis described in §§ III (7) and V (2). Occasionally the same concentration of hydroxylamine brings the reaction down to the incomplete form of reaction (*d*). The results obtained when hydroxylamine is added aerobically *before* adaptation parallel those observed with photoreduction. This is not surprising because hydroxylamine in rather small concentrations interferes with the adaptation reaction, and complete adaptation is necessary for the oxyhydrogen reaction as well as for photoreduction.

In all the observations mentioned up to now we see a similarity between photoreduction and dark reduction of carbon dioxide in respect to the influence of cyanide and hydroxylamine. Now a significant difference has to be pointed out, which concerns the reversion to aerobic conditions. The reversion brought about by a higher light intensity is retarded or prevented by hydroxylamine. No such 'protection' by hydroxylamine exists if the reversion is enforced by an excess of oxygen. With an excess of oxygen the hydrogenase becomes inactive whether hydroxylamine is present or not.

(4) *Partial reactions.* Since the experiments show definitely that the oxyhydrogen reaction consists of a sequence of partial reactions we can ask two pertinent questions: (1) which partial reaction is coupled with the reduction of carbon dioxide? and (2) which partial reaction leads to inactivation if its rate becomes too fast? In answer to the first question we can state that in absence of carbon dioxide or in presence of inhibitors which prevent the reduction of carbon dioxide the oxyhydrogen reaction proceeds to the 'peroxide' level with the utilization of only one molecule of hydrogen. It is the second molecule of hydrogen, therefore, whose absorption enables a third molecule of hydrogen to be transferred to carbon dioxide (or its equivalent, a carbon dioxide complex, carboxyl group, etc.).

Contrary to photoreduction, where an inactivation by light takes place only in presence of carbon dioxide, the inactivation of the hydrogenase system by oxygen is quite independent of the available carbon dioxide. This is obvious, because in the first case the oxidizing intermediates are formed in the course of the photochemical reduction of carbon dioxide, whereas in the second case they are formed directly from oxygen. The experiments with cyanide, on the other hand, have shown that an inactivation is practically inevitable in presence of this poison, even if reaction (*d*) leading to the peroxide level is not inhibited at first. This agrees well with the corresponding results on photoreduction in presence of cyanide. Again, it seems that a peroxide intermediate causes inactivation whenever it is allowed to accumulate. We assume that this intermediate is very similar to the substance which produces the inactivation in excess light. This assumption, however, is arbitrary and offers at the moment only an economy of symbols in the theoretical treatment of the problems involved (§ VII).

(5) *Oxyhydrogen reaction and respiration.* In absence of carbon dioxide little more than one molecule of hydrogen is absorbed for each molecule of oxygen. The oxyhydrogen reaction proceeds apparently to the peroxide level only. If under these conditions the 'peroxidic' intermediate would accumulate, the reaction should soon come to an end because the hydrogenase system should be progressively inactivated. Actually the incomplete oxyhydrogen reaction has been repeated with 0·5 % of an atmosphere up to ten times in succession with the same sample of cells without a decrease in activity. This may be due to two different causes. Either one-half of the oxygen does not enter into the oxyhydrogen reaction but is used up by respiration while the other half forms water, or all the oxygen forms the 'peroxidic' intermediate, and this in turn is reduced immediately after its formation by intracellular hydrogen donors. In order to decide between these possibilities the total amount of carbon dioxide formed during the incomplete oxyhydrogen reaction has been determined. It has been found that the increase in carbon dioxide amounts to less than one-seventh of the quantity which would be formed if half of the oxygen would be used up by normal respiration. The amount is even smaller than that produced in fermenting control cells. These experiments tell us that the added oxygen is reduced faster by the hydrogenase than by the ordinary respiratory system. It follows that all oxygen is first reduced to the 'peroxide' and that the latter disappears by way of an internal reduction which yields no carbon dioxide. The assumption of an internal reduction is further supported by the fact that in these algae the hydrogenase establishes an equilibrium between external and internal hydrogen donors. The potential of the various hydrogen donors and the partial pressure of hydrogen determines the direction of the hydrogen transfer

(§ II (2)). The interesting aspect of this reaction is that the presence of carbon dioxide determines whether the intermediate 'peroxide' is reduced faster by molecular hydrogen or by internal hydrogen donors.

We have seen above (§ III (6)) that addition of glucose to the algal suspension decreases the rate at which hydrogen is used in photoreduction. If the explanation given for this effect is correct, namely, if an increased transfer of hydrogen from carbohydrate derivatives competes with external hydrogen, the same effect of glucose could be expected to appear also in the oxyhydrogen reaction. A few preliminary experiments have shown indeed that much less hydrogen is absorbed (even in presence of carbon dioxide) after the algae have been fed with glucose.

As mentioned in § VI (3), the 'incomplete' oxyhydrogen reaction entails no harm to the reduced active state of the hydrogenase, but leads to its inactivation, if cyanide is present. We must assume, then, that the internal reduction of the peroxidic intermediate is inhibited by cyanide in a similar way as its reduction by the second molecule of external hydrogen.

VII. DIAGRAMS DESCRIBING THE REACTIONS INVOLVING CARBON DIOXIDE IN PHOTOSYNTHESIZING CELLS

In the past no other biological problem has been more conducive, for lack of true knowledge, to purely hypothetical speculation than photosynthesis. The increasing number of experimental data which have become available during recent years, however, has made all the earlier detailed theories of the chemical mechanism of photosynthesis untenable. In fact only the very general formulation of the problem as presented in equation (*f*)

$$n\,[CO_2 + 4H_2O + 10h\nu] \rightarrow (CH_2O)_n + 3n\,[H_2O] + nO_2, \quad (f)$$

is not contradicted by the experimental results. The most explicit interpretation concerning the properties of the photochemical process is given in a recent theoretical paper by Franck & Herzfeld (1941), the outcome of several years of discussion (Briggs, 1941; Gaffron & Wohl, 1936; Ornstein, Wassink, Reman & Vermeulen, 1938; Wohl, 1937, 1941). The observations on the anerobic metabolism of algae would add little to a better understanding of the process of photosynthesis if they could not be correlated with each other, with the older experiment and with the latest theoretical conclusions. We shall, therefore, summarize the material presented in this article by drawing a comprehensive diagram demonstrating the possible interrelation of the various reactions occurring during the reduction of carbon dioxide with water or hydrogen. First we have to interpret the single observations, then we have to translate the interpretations into symbolic chemical equations, finally we construct a general diagram based on these equations.

(1) *Explanation of the multiple effects of specific poisons.* In low concentrations hydroxylamine inhibits photosynthesis. Since these amounts do not interfere either with photoreduction or with the reversion reaction we assume that only the liberation of oxygen is involved. The same concentrations, however, interfere with the adaptation reaction. Hence we conclude that, somehow, the oxygen-liberating system is connected with the process of adaptation. After adaptation large amounts of hydroxylamine lower the rate of photoreduction without ever stopping it completely and also prevent the reversion reaction up to very high light intensities. Since the effect upon the rate of photoreduction is noticeable at rather low light intensities it is probable that the photochemical reaction is slowed down. The question arises whether this would also automatically explain the protection against a reversion to photosynthesis. The reversion is brought about by an excess of 'photoperoxides' (see § V (1)), and it may be prevented in two ways: Either by diminishing the rate of the photochemical process so that no excess of peroxides is formed or by preventing an actual surplus of peroxides from interfering with the activity of the hydrogenase system. The first case is exemplified by the stability of the reduced state at high light intensities in the absence of carbon dioxide or in presence of carbon dioxide and dinitrophenol (§ V (6)). Here the rate of photoreduction is kept so low that it cannot exceed the reversion threshold. With hydroxylamine, on the other hand, we obtain the highest rate of photoreduction possible (§ V (3)). Here, then, the second case obtains, a direct inhibition of the reversion reaction. Since the reversion includes the inactivation of the hydrogenase system, and the latter is brought about by the peroxides, it follows that photoreduction does not depend on the formation of 'peroxides'. It is unlikely, therefore, that the peroxides are the intermediate products arising directly from the photochemical decomposition of water. A first intermediate step must be inserted which is followed by the formation of peroxide as a second step. And we are forced to assume that it is to this intermediate, the primary oxidized product, that the hydrogenase system transfers hydrogen. As to the hydroxylamine poisoning, we still have one choice to make. We can assume either that the peroxides combine with hydroxylamine and in this form do not react with the hydrogenase, or that the formation of the peroxides from the primary oxidized products is inhibited. The first alternative would also explain the aerobic effect of hydroxylamine on photosynthesis, but it would not explain why a nearly ten times higher concentration of hydroxylamine is needed to protect photoreduction (and that incompletely) than is necessary for a complete inhi-

bition of the oxygen production. We therefore prefer the second explanation. It leads to the consequence that the evolution of oxygen from the peroxide is inhibited by small amounts of hydroxylamine, independently of the hydroxylamine poisoning of the reaction by which the peroxides are formed.

If we translate the sequence of reactions required by our explanation into chemical symbols, we obtain the equations (1)–(9). Besides H_2, O_2, HOH and CO_2 the symbols used have the following meaning. X and Y are a pair of catalysts which make it possible for water to be decomposed by the small quanta of red light. Normally they form a complex with carbon dioxide (Ruben, Kamen & Hassid, 1939), water (Pratt & Trelease, 1938) and chlorophyll. Z stands for an oxygen transporting intermediate. The hydrogenase system appears in three forms: HDo, active, hydrogenated; Do, active, dehydrogenated; DoOH, inactive, oxidized. GH means any other internal hydrogen donor except H_2. Fr. signifies an enzyme which liberates free oxygen from a precursor (peroxide or moloxide). Among these reactions the formation of the peroxide (equation (2)) is inhibited if hydroxylamine reacts with the auxiliary substance Z; the formation of oxygen if hydroxylamine combines with the peroxide-decomposing catalyst Fr.

The fact that when reaction (2) is blocked photoreduction continues at light intensities which must produce an excess of YOH shows clearly that YOH is not a dangerous substance; it cannot inactivate the hydrogenase Do even if reaction (3) proceeds at such a rate that the stationary concentration of the dehydrogenated hydrogenase Do is high. The fate of the unused oxidized photoproducts YOH is most likely a back reaction (equation (8)) with the reduced counterpart $CO_2 \cdot XH$ formed in equal amounts by the photochemical decomposition of water according to equation (1). Such a concept explains why any interference on the oxygen side of the mechanism automatically diminishes the photochemical yield of the reduction of carbon dioxide and makes it understandable why no accumulation of reducing substances has ever been observed under the influence of poisons. The inactivation reaction (6) has, of course, its counterpart in the activation reaction (7). Reactions (6) and (7) together yield

$$Z(OH)_2 + 2HDo \rightarrow Z + 2Do + 2H_2O.$$

This means that photoreduction may also proceed undisturbedly via the peroxides, so long as there is an excess of reduced hydrogenase. The experiments (§ V (1)) tell us that the reactivation (7) is rapid if the inactivation (6) has been incomplete, but that (7) requires a long period of fermentation if all of the hydrogenase system Do has been oxidized. Reaction (7) explains this observation by showing that active molecules can 'rescue' the inactivated ones. The activation process is autocatalytic and must have a long induction period if all molecules of the hydrogenase have been oxidized. The activation then depends on a side reaction with GH supplying the first molecule Do capable of transferring hydrogen.

The question which now arises is whether the various effects of cyanide contradict any of the equations discussed above. Experimentally we can distinguish three reactions which are influenced by cyanide: the fixation of carbon dioxide, the adaptation to hydrogen and the reversion. We can explain why cyanide accelerates the reversion if we remember that the adaptation can be prevented by very small concentrations of cyanide. Describing the latter

Photochemical reaction: $CO_2 \cdot X \cdot Y : HOH \xrightarrow{\text{light}} XH \cdot (CO_2) + YOH$, (1)
(photosensitive complex; reduced product; oxidized product)

Formation of peroxide: $2YOH + Z \rightarrow Z(OH)_2 + Y$, (2)
($2YOH$: first oxidized product; Z: auxiliary compound (enzyme and substrate); $Z(OH)_2$: 'peroxide')

Reduction by reduced hydrogenase: $YOH + HDo \rightarrow H_2O + Y + Do$, (3)

Uptake of hydrogen by oxidized hydrogenase: $Do + \frac{1}{2}H_2 \rightarrow HDo$, (4)

Liberation of oxygen: $2Z(OH)_2 + Fr. \rightarrow Fr.O_2 + 2Z + 2H_2O \rightarrow Fr. + O_2$, (5)

Inactivation of hydrogenase: $Z(OH)_2 + HDo \rightarrow Z + H_2O + DoOH$, (6)

Reactivation of hydrogenase: $DoOH + HDo \rightarrow 2Do + H_2O$ (7)

or: $DoOH + GH \rightarrow Do + G + H_2O$, (7a)

Back reaction to (1): $X \cdot H \cdot (CO_2) + YOH \rightarrow CO_2 \cdot X \cdot Y + HOH$, (8)

Same in absence of carbon dioxide: $XH + YOH \rightarrow X \cdot Y + HOH$, (8a)

Reversible change of the oxygen-liberating enzyme: Fr. (Cat.) $\rightleftharpoons$ Fr. (Peroxy.). (9)
(catalase activity; peroxydase activity)

effect (§ II (3)) we said that in all probability cyanide combines with the oxidized inactive form of the hydrogenase system and stabilizes it against reduction. If we assume a completely reduced, that is active, hydrogenase system and add cyanide, the poison, ineffective at first, will trap all the molecules of the catalyst which happen to be oxidized to DoOH in the course of time. In the unpoisoned stationary state of photoreduction the presence of inactive molecules is not noticeable because their concentration remains small (velocity of (7) equal to that of (6)). If, however, we make the oxidation irreversible by adding cyanide, the eventual reversion to photosynthesis is inevitable. With hydroxylamine we can retard the irreversible destruction of the hydrogenase activity by decreasing the rate at which peroxide molecules are formed. The chances for an oxidation

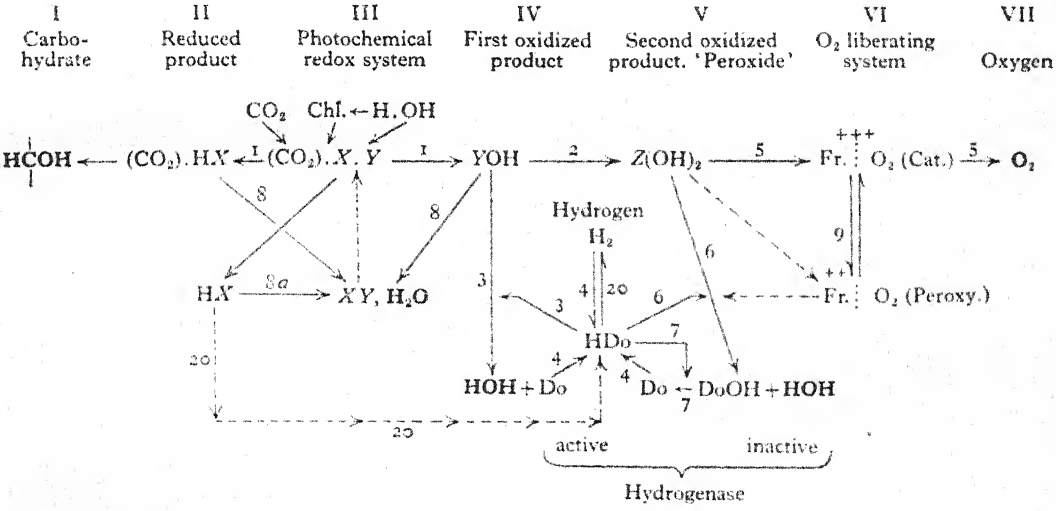


Fig. 1. Diagram demonstrating the relation between photosynthesis and photoreduction.

might think that the adaptation reaction could be described adequately by the reactions (7*a*) and (7), a process which seems to proceed very slowly until the first activated hydrogenase molecules begin to participate in the hydrogen transfer. After that the adaptation is autocatalytically accelerated. As to the effects of poisons, we can assume that carbon monoxide inhibits all reactions of the reduced active hydrogenase, thereby accelerating the inactivation. Cyanide apparently combines with the hydrogenase in its oxidized state and prevents its reduction. But there is no experimental indication that small concentrations of hydroxylamine interfere with the reactions of either form of the hydrogenase systems. Yet small amounts of hydroxylamine inhibit adaptation. This indicates that more than one enzyme is involved in the adaptation reaction. The effect of

of the hydrogenase molecules are thus diminished, but we cannot prevent it, at least so long as the protection given by hydroxylamine is not complete.

We see that the equations (1)–(7) describing the pathway of the oxidized photoproduct, which were written to fit the hydroxylamine experiments, agree well with the observed action of cyanide, except that the latter emphasizes the following point. The number of peroxide molecules formed in the normal course of photoreduction cannot be negligible if our explanation of the cyanide effect is to hold. In other words, during photoreduction the flow of oxygen to hydrogen is divided; one branch goes directly from YOH to Do (equation (3)), while the other branch goes via the peroxides (equations (2), (6) and (7)).

Another observation which must be discussed is the inhibition of the adaptation to hydrogen by the same small concentrations of hydroxylamine which stop the evolution of oxygen (§ II (3)). At first one

hydroxylamine on the evolution of oxygen suggests that in addition to the hydrogenase the oxygen-liberating system also undergoes a change during the adaptation period (reaction (9)). For this hypothesis we have, however, as yet no further experimental support. We can assume that the oxygen-liberating system, whose function is analogous to that of *catalase*, acquires during the adaptation the capacities of a *peroxidase* or of an *oxidase* and serves as such in the hydrogenase cycle. In the vast literature concerning the properties of iron-porphyrin catalysis one occasionally finds a hint that such changes *are* possible (Swedin & Theorell, 1940; Theorell, 1940; Keilin & Hartree, 1935–6). The postulated change is not in contradiction with the fact that the high affinity for hydroxylamine disappears after adaptation. Most oxidases and peroxidases are not inhibited *in vivo* by hydroxylamine in concentrations which paralyse the activity of catalase (an interesting

exception has been found in *Azotobacter* (Wilson & Wilson, 1943)).

(2) *The relation between photosynthesis and photoreduction.* The equations (1)–(9) lead us to the following scheme (Fig. 1) covering the photochemical processes described in this article. Fig. 1 consists of seven sections representing the major subdivisions of the process of photosynthesis. The choice of seven sections is a compromise between the desire to avoid purely hypothetical assumptions and the attempt to correlate all observations. Several of these sections will undoubtedly be further subdivided in the future, others may be omitted in a more compact scheme. The arrows indicate the progress of the different partial reactions from one intermediate step to another in the direction of the end products: carbohydrate, oxygen and water. The arrows do not represent balanced reactions, but are given the numbers of the corresponding complete equations. It is clear, for example, that for each molecule of water entering in (iii) only one fourth molecule of oxygen can appear in (vii). We shall proceed now to discuss the diagram section by section.

(i) *Carbohydrate.* In the stationary state, after all induction periods have passed, the assimilatory quotient $AQ = +\Delta O_2/-\Delta CO_2$ is -1 in photosynthesis. In photoreduction two volumes of hydrogen are absorbed per one volume of carbon dioxide and the quotient $AQ = -\Delta H_2/-\Delta CO_2$ is 2. This has been taken as a proof that the reaction product in both reactions is a carbohydrate or a $-\overset{|}{C}HOH$ group in a carbon chain. The nature of the first carbohydrate is unknown. For convenience we assume that the same first product is formed in all three synthetic processes (photosynthesis, photoreduction and dark-reduction or chemo-synthesis).

(ii) *Reduced product.* Four hydrogen atoms have to be transferred to carbonic acid or its organic derivative until the carbohydrate level has been reached. Hence the intermediates must pass several times through the oxido-reduction cycle (iii) before a new carbon dioxide molecule is added to the chlorophyll complex. An alternative has been suggested by Rabinowitch, namely, a dismutation of four primary reduced molecules to one carbohydrate and three carbon dioxide molecules:

$$4H(CO_2) \rightarrow [H_4CO_2] + 3CO_2.$$

Dinitrophenol inhibits photosynthesis and photoreduction and enhances the photochemical production of hydrogen—an effect similar to that caused by the absence of carbon dioxide. It is, therefore, probable that the transfer of hydrogen involved in the formation of the reduced product is the reaction sensitive to this poison. In absence of carbon dioxide or in presence of dinitrophenol only the short-lived compound XH is formed instead of the reduction product $X \cdot H \cdot CO_2$. XH reacts back rapidly with its oxidized partner YOH. If the algae are adapted to hydrogen some YOH may be reduced by hydrogen donors other than XH. Thus some XH is left over and may transfer its hydrogen to the hydrogenase Do (arrow 20) from where it is released in molecular form. This may be the mechanism of the photochemical hydrogen production (§ IV (1)).

(iii) *Photochemical redox system.* The function of the photochemical apparatus is probably the same in photoreduction as in photosynthesis. This statement is based mainly on two facts. First, the quantum yield is identical in both reactions, though the energy requirements for the over-all reaction differ by 100,000 cal. Secondly, the flash saturation is also the same for both reactions. We picture the substrate of the photochemical reaction as a complex containing carbon dioxide, water, a catalyst (organic) XY, and chlorophyll. Whether chlorophyll merely sensitizes the hydrogen transfer in other parts of the complex or participates directly in this transfer is a question which is not essential for the present scheme. Franck & Herzfeld (1941) have presented some reasons why the transfer of a hydrogen atom from water to carbon dioxide with the aid of two quanta may include chlorophyll as an intermediate hydrogen acceptor.

The first reduced intermediate XH must have a short lifetime because experiments have shown that fixed carbon dioxide has to be present during the absorption of light, otherwise the energy is lost for photosynthesis (Emerson & Arnold, 1932). (Compare the need of carbon dioxide for the oxyhydrogen reaction, § VI (2).) The back reaction (8) apparently prevents any accumulation of reducing substances which subsequently reduce carbon dioxide slowly in the dark.

The reduction of carbon dioxide is quite specific; no organic substance, so far, has been found which might replace it. We know further that the photochemical process is preceded by a reversible reaction of carbon dioxide with an unknown large molecule (Ruben *et al.* 1939, 1940). In the scheme the catalyst which combines with carbon dioxide and water is symbolized by XY. Under the influence of light this compound splits into a reduced part HX and an oxidized part YOH. Only inasmuch as YOH is used up (by reactions which lead either to oxygen or to water) is the part HX at liberty to transfer its hydrogen to carbon dioxide. Otherwise we have to assume a recombination of XH and YOH with the formation of water. This is a way by which light energy may be lost under conditions which are unfavourable for its chemical utilization (poisoning with cyanide, dinitrophenol, urethane, etc.).

(iv) *Oxidized product.* The effect of hydroxylamine in preventing the reversion from photoreduction to photosynthesis has led us to assume the existence of a first oxidized product, an intermediate between the photochemical redox system and the 'peroxide'. Since the reaction (2) must be spontaneous and fast it is clear that $2\,YOH$ have to have a larger free-

energy content than $Z(OH)_2$. The mechanism of reaction (2) remains quite obscure. The formation of the 'peroxide' is the next step in photosynthesis under aerobic conditions, since the side roads to the hydrogenase system in (iv) and (v) remain closed as long as the latter system is kept inactive by the presence of oxygen, i.e. as long as all the hydrogenase molecules are in the oxidized state DoOH. After the adaptation to hydrogen the hydrogenase is in the form Do which is able to react with hydrogen, forming HDo. HDo and *Y*OH yield water and Do and *Y* become free for a new cycle. The easy reversion from this process to the evolution of oxygen shows that the transition from the oxidized product to the 'peroxide' is not inhibited during photoreduction: water is formed in place of oxygen merely because reactions (6), (7) and (4) are faster than reaction (5).

(v) '*Peroxide*'. As said above, the effect of cyanide upon photoreduction can be understood only if we assume that the 'peroxide' is always formed in adapted algae. The peroxide formed despite reaction (3) disappears by the second side road leading to the hydrogenase system and the formation of water (reactions (6), (7)). The latter reaction, however, is comparatively slow. This will not be noticed so long as there remains an excess of active (hydrogenated) hydrogenase because of the reactions (7) and (4), for instance at low light intensities. The more *Y*OH is produced by increasing illumination, the less HDo is available to reactivate the oxidized hydrogenase molecules. At the same time more and more of the inactivating peroxide is formed. At a certain intensity this autocatalytic process must enforce a return to an aerobic metabolism. Cyanide accelerates the inactivation of the hydrogenase system because it prevents any reactivation. Hydroxylamine retards inactivation because it inhibits the reaction (2) leading from the oxidized product to the peroxide. At high light intensities, any excess of *Y*OH disappears by back reaction (8), (8*a*).

(vi) *Oxygen-liberating system.* The peroxide (v) must be converted into oxygen by a specific enzyme. This assumption is required in order to explain the effect of *small* amounts of hydroxylamine, which poison the liberation of oxygen, but not the absorption of hydrogen after adaptation to photoreduction (§ III (7)). Concentrations of hydroxylamine which suffice to inhibit photosynthesis are insufficient to protect the hydrogenase system against photochemical inactivation. Under these conditions a strong illumination leads to a complete cessation of all gas exchange. Because the adaptation is inhibited by hydroxylamine in the same manner as the liberation of oxygen, we introduced reaction (9). The oxygen-liberating enzyme is supposed to change into a form in which it is capable of transferring oxygen either from the peroxide or from the air to the hydrogenase system. This hypothetical conversion may be brought about by a valency change in a prosthetic group or by an exchange of protein (Gjessing & Sumner, 1942).

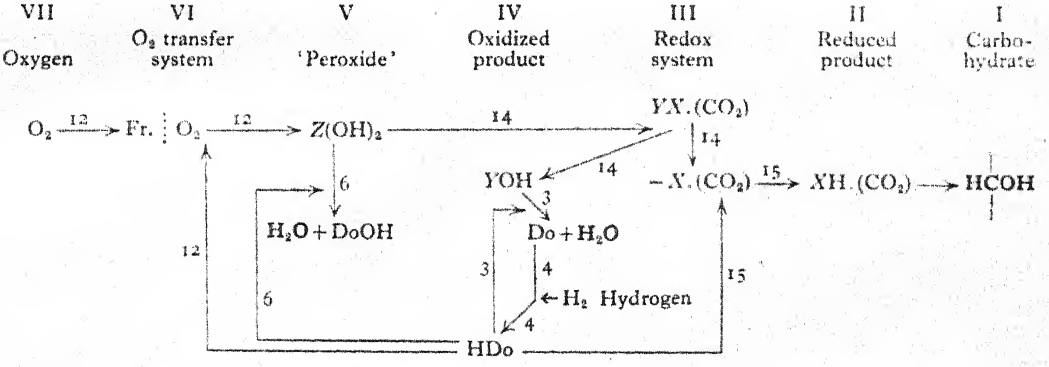


Fig. 2. Diagram demonstrating a coupling between oxyhydrogen reaction and carbon dioxide reduction.

(vii) *Oxygen.* Warburg found a depressing influence of oxygen on photosynthesis (Warburg, 1928). This effect varies greatly with the light intensity and with nutritional conditions. It may be indirectly connected with the photoxidation which occurs at very high light intensities (Burr & Myers, 1940) or in the absence of carbon dioxide (Franck & French, 1941). The photoxidation is not influenced by specific poisons. Perhaps the reduced product *X*H is oxidized directly. After adaptation to the reduced state the influence of oxygen changes completely. Now the photochemical mechanism is connected with an oxygen-transferring system, and we observe a direct competition of oxygen with the course of the light reaction. This we shall discuss in connexion with Fig. 2 explaining the oxyhydrogen reaction.

(3) *The coupling of the oxyhydrogen reaction with the reduction of carbon dioxide.* The present knowledge of the respiratory mechanism in living cells enables us to draw a simple scheme explaining the

formation of water from its elements *in vivo*. We need only to combine a hydrogen transfer system with an oxygen transfer system, i.e. a hydrogenase with an oxydase. This combination has been proposed for instance by Yamagata & Nakamura (1938) to explain the oxyhydrogen reaction in purple bacteria. In washed suspensions of luminous bacteria Claren has even demonstrated the catalytic role of a fumarate-succinate system in the oxyhydrogen reaction (Claren, 1938). Yet we would like to understand not only the way by which water is formed, but also the way by which this mechanism is coupled with the reduction of carbon dioxide. Ruhland (1924) found a very efficient coupling between the oxyhydrogen reaction and the reduction of carbon dioxide in bacteria. Our experiments have shown the same phenomenon in the adapted algae. The problem is: How can carbon dioxide be involved in the oxyhydrogen reaction? The following scheme (Fig. 2) is an attempt to answer this question. The main points on which it is based, are: (1) the quantitative relation expressed by the equation (*c*):

$$CO_2+2O_2+6H_2 \rightarrow (CH_2O)+5H_2O; \quad (c)$$

(2) the observation that the two hydrogen molecules needed for the complete reduction of water are not equivalent in respect of their relation to the coupled reduction; and (3) that it often depends on the presence of carbon dioxide whether the second molecule of hydrogen is absorbed at all. Furthermore, the general conditions favourable to the oxyhydrogen reaction are identical with those which support photoreduction; and a great similarity exists between the two metabolic processes in regard to the action of poisons. We start therefore with the assumption that some or even all the catalytic systems involved in photoreduction are also used in the oxyhydrogen reaction. For convenience we again divide the diagram into seven sections, this time reversing the sequence of events. The process starts with the absorption of oxygen instead of ending with its liberation.

(vii), (vi), (v) *Oxygen and peroxide*. The peroxide is the first intermediate for the existence of which we have good evidence. Its formation from the elements may be written according to reactions (10) and (11):

$$O_2+Z \rightarrow ZO_2, \quad (10)$$

$$ZO_2+2HDo \rightarrow Z(OH)_2+2Do. \quad (11)$$

(vi), the oxygen-transferring system, is not demanded directly by experiment. It is unlikely, however, that the 'peroxide' can be formed without the aid of an autoxidizable oxidase. We prefer therefore reaction (12):

$$O_2+Fr. \rightarrow FrO_2;\ FrO_2+Z+2HDo \rightarrow Fr.+2Do+Z(OH)_2. \quad (12)$$

In absence of carbon dioxide the absorption of hydrogen ends after the transfer of one molecule of hydrogen. The peroxide is then reduced by some internal hydrogen donor GH (reaction (13)):

$$Z(OH)_2+2GH=Z+2G+2H_2O. \quad (13)$$

Cyanide does not stop this reaction immediately, but causes a rapidly increasing deactivation of the whole hydrogenase system ending in a return to aerobic conditions. This indicates that the same cycle of active and inactive hydrogenase which is shown in (v) of the first scheme above must occur here. Wieland & Pistor (1936, 1938) are convinced that the oxyhydrogen reaction in living cells must proceed with hydrogen peroxide as an intermediate. They arrive at this conclusion because in *Acetobacter peroxidans* hydrogen peroxide seems to be reduced faster than molecular oxygen. We have added hydrogen peroxide to a suspension of adapted algae and have found that it is decomposed much faster by catalase (anaerobically! see Sumner, 1941) than it is reduced by hydrogenase. Up to a concentration of $10^{-3}M$ hydrogen peroxide the hydrogenase remained active. In presence of an initial hydrogen peroxide concentration of $2\times10^{-3}M$ we observed a reversion to aerobic conditions. This may have been caused just as well by oxygen accumulating in the suspension as by the hydrogen peroxide itself. The assumption that hydrogen peroxide must be an intermediate in the oxyhydrogen reaction does not hold in our case. The energy released in the decomposition of hydrogen peroxide by catalase is lost for any activation of the carbon dioxide reduction. Furthermore, if hydrogen peroxide and catalase should participate, at least some influence of hydroxylamine in small concentrations upon the course of the oxyhydrogen reaction would be noticeable, because catalase is so readily inhibited by this poison. In short, the available data on the inactivation reaction are most conveniently explained as due to an excess of an intermediate at the peroxide oxidation level, which is distinctly different from hydrogen peroxide. The excess may be caused either by too much oxygen, so that the intermediate peroxide is formed faster than the cell is able to reduce it, or by an inhibition of the hydrogen transfer to the peroxide.

(iii), (iv) *Redox system and oxidized product*. The next step, the formation of water from 'peroxide' and hydrogen proceeds smoothly only in the presence of carbon dioxide and is coupled with the reduction of the latter. The stoichiometric relations must be such that equal amounts of hydrogen react with the peroxide on the one hand and with carbon dioxide on the other:

$$2H_2+2Z(OH)_2 \rightarrow 2Z+4H_2O;$$
$$2H_2+XCO_2 \rightarrow X(H_2CO_2)+H_2O.$$

We interpret this coupled reaction as a breaking apart by oxidation of a complex $Y\cdot X(CO_2)$ into an oxidized compound YOH and a radical $-X(CO_2)$.

The latter receives hydrogen from the hydrogenase and forms a reduced product $HX(CO_2)$.

$$Z(OH)_2 + 2\,Y\cdot X(CO_2) \rightarrow Z + 2\,YOH + 2\text{—}X(CO_2), \quad (14)$$

$$2\text{—}X(CO_2) + 2HDo \rightarrow 2X(H)\,(CO_2) + 2Do, \quad (15)$$

$$2\,YOH + 2HDo \rightarrow 2\,Y + 2Do + 2H_2O, \text{ etc.} \quad (3)=(16)$$

The oxidized product is finally reduced to water. In other words, we produce by an oxido-reduction in the dark the same (or quite similar) substances as those which appear in the photochemical reaction.* These substances react further as described in the scheme representing the light reaction. When the series of cyclic reactions have run out two molecules of hydrogen have been transferred to oxygen and one to carbon dioxide.

The experimental observations require us to assume that the redox system XY functions only if carbon dioxide is bound to it. In cases when carbon dioxide is not available or when its fixation is prevented by cyanide, only reactions (12) and (13) occur. It happens sometimes that the oxyhydrogen reaction goes to completion and two molecules of hydrogen are absorbed, yet no carbon dioxide is reduced. This we can interpret as taking place with the aid of the complex $Y.X.CO_2$ under conditions where the transfer of hydrogen inside the complex to the carbon dioxide derivative is inhibited. Reactions (14) and (15) are followed by a back reaction (17) but not by the formation of a stable reduced product:

$$X(H)\,(CO)_2 + YOH \rightarrow X\cdot Y\cdot CO_2 + H_2O. \quad (17)$$

We do not believe that Fig. 2 describes the only way by which the oxyhydrogen reaction could induce a reduction of carbon dioxide; but it is important to note that if one tries to correlate the quantitative data on the dark reduction of carbon dioxide in *Scenedesmus* with the effects of different poisons, one is led to a scheme which is practically the reverse of that designed for the interpretation of the photochemical processes.

(4) *Remarks on the validity of the diagrams presented.* The schemes have been put together so as to describe the behaviour of algae under anaerobic 'adapted' conditions. To be of general use, they must not contradict any essential data on photosynthesis. The diagrams are incomplete with regard to the dark fixation of carbon dioxide, but there are no difficulties in extending (iii) so as to include the fixation process (18)

$$RH + CO_2 \rightarrow RH\cdot CO_2; \quad RH\cdot CO_2 \xrightarrow[-A]{+A} R\cdot COOH, \quad (18)$$

where A is a cyanide-sensitive catalyst (Ruben *et al.* 1939, 1940; van Niel, Ruben, Carson, Kamen & Foster, 1942).

If we compare the diagrams with van Niel's (1941) latest description and interpretation of the metabolism of the purple bacteria, we find a perfect agreement. This is no wonder, since the adapted algae utilize hydrogen exactly like certain purple bacteria (Roelofsen, 1935; Gaffron, 1935*a*). The phenomena are the same in both classes of organisms except for the transition reactions. Purple bacteria need no long anaerobic adaptation period before their hydrogenase becomes active, nor do they ever produce oxygen. This difference offers a hint that the hypothetical change of the oxygen-liberating system, reaction (9), may be the decisive factor in the adaptation of green algae.

A large amount of experimental work and interpretative ingenuity has been devoted to the induction phenomena in photosynthesis (Aufdemgarten, 1939; McAlister, 1939; McAlister & Myers, 1939; Gaffron, 1940*b*), by which is meant all kinds of deviations observed during the first moments of illumination. Some years ago it was suggested that a part of the oxygen-liberating system could change reversibly from an active (reduced) form to an inactive (oxidized) one (Gaffron, 1937). The oxygen of the air was supposed to inactivate most of the enzyme in the dark and the reduced products of photosynthesis to reactivate it in the light. This hypothesis explained the initial but passing inhibition after a dark period. Later experiences showed that this idea, after some modification and extensions (Franck *et al.* 1941), can explain a great variety of observations on the induction period which have recently been made with the aid of fast recording devices. Fig. 1 requires that for an oxygen evolution at the normal rate the distribution of the reduced and oxidized forms of all catalysts, Y, Z, Fr., has to be that of the stationary state in continuous illumination. If, after a dark period, one or several of the carriers are present mainly in the oxidized form, an induction period is unavoidable. The diagram also indicates that we may expect different induction phenomena for the release of oxygen and for the absorption of carbon dioxide (Franck & Herzfeld, 1941).

A conspicuous difference between the reduction of carbon dioxide in the light and in the dark is the course of both reactions in presence of the higher concentrations of hydroxylamine. The rate of photoreduction at moderate intensities is clearly inhibited, whereas the rate of the oxyhydrogen reaction remains generally unchanged. Simultaneously we find that

* Dr Rabinowitch, to whom this scheme was submitted for criticism, believes it would be better to write:

$$YOH + \mathbf{H_2Do} \rightarrow Y + H_2O + HDo-, \quad (16a)$$

$$HDo- + X(CO_2) \rightarrow HX(CO_2) + Do. \quad (15a)$$

He assumes that the oxidation does not produce a radical $-XCO_2$ from the complex $XY(CO_2)$ but is followed by reaction (16*a*) yielding a highly reactive, half-oxidized hydrogen donor HDo−. If one prefers this version, one has to give to the fully reduced hydrogenase the formula H_2Do and change the diagrams accordingly.

the anaerobic metabolism is protected against the reversion caused by light but not against that caused by oxygen. Comparing the two schemes we see that this difference may result from the fact that reaction (2) above has no place in the scheme describing the oxyhydrogen process. And it is probably reaction (2), as already pointed out, which is inhibited by hydroxylamine. Why only reaction (2) and not reaction (14) should be sensitive to hydroxylamine, why in some cases the same amount of hydroxylamine destroys the coupling between the (complete) oxyhydrogen reaction and the reduction of carbon dioxide and in other cases not, why cyanide inhibits the coupled reduction of carbon dioxide more promptly than the formation of the 'peroxide', are but a few of the questions which the scheme helps to formulate, but in its present state is not good enough to answer. It should be pointed out that such diagrams (and even more so the equations (1)–(16)) can be misleading by giving the impression as if all partial reactions occurred between dissolved molecules moving freely in a homogeneous medium. Actually the binding of substrates, intermediates and products to the surfaces of large protein complexes may result in a much stronger coupling between the single reaction steps.

The question whether the oxidation of a complex $CO_2 \cdot XY$ can yield enough energy for the transfer of hydrogen to carbon dioxide may be answered by pointing out that the oxidation of the luciferin-luciferase complex in *Cypridina* or in luminous bacteria leads to the emission of visible light quanta (Harvey, 1940).

If the way of the dark reduction corresponds with that of the photochemical reduction of carbon dioxide we can expect a strong interference between both processes.* In the algae it has been established that the maximum rates of hydrogen absorption which are tolerated without a reversion are equal for both processes. Light and oxygen may be applied simultaneously so long as the utilization of hydrogen does not surpass the critical rate; in other words both compete for hydrogen from the same donors. More precise experiments are made difficult by the easy reversion to aerobic conditions.

Fortunately, the literature on purple bacteria offers a good example of interference between oxidative and reductive reactions as required by our schemes. It had been observed that light decreases the rate of oxygen absorption by purple bacteria (Gaffron, 1935*a*; Nakamura, 1937). From all the other available evidence it was clear that this could not be interpreted as a photochemical production of free oxygen by the bacteria. Rather one had to assume a competition of the light and the dark reactions for the same hydrogen donor (Gaffron, 1937; Wassink, Vermeulen, Reman & Katz, 1938), and hence the existence of an oxygen-transferring system in close interrelation with the photochemical mechanism. Recently clear-cut quantitative experiments by van Niel (1941), on the metabolism of *Spirillum rubrum*, a purple bacterium, have shown that oxygen on the one hand and carbon dioxide (plus light energy) on the other can replace each other as hydrogen acceptors. In light, added acetate is consumed solely in photoreduction despite the presence of oxygen. In the dark the same amount of acetate is oxidized at the same rate by way of respiration. True, acetate is not molecular hydrogen, and the hydrogenase of the algae has different properties from the acetodehydrogenase of a purple bacterium. The latter, for instance, is not inactivated by an excess of oxygen. But it is obvious that with slight modifications our diagrams can be applied also to the reactions in *Spirillum rubrum*. And the similarity becomes even more striking when we read (van Niel, 1941, p. 305) that 'the oxidative decomposition [of acetate] is apparently dependent on the presence of carbon dioxide' (cf. Hes, 1938 *a*, *b*).

Summing up we can say that the diagrams resulting from equations (1)–(15) are not contradicted by any of the earlier observations on photosynthesis. They underline once more the validity of the general principle of oxido-reduction and hence fully support the arguments presented in two earlier reviews (Franck & Gaffron, 1941; van Niel, 1941). They constitute a progress in so far as they propose a definite mechanism for a series of partial reactions which have become distinguishable by the experiments reviewed in this article. Finally, the diagrams are a challenge to the experimenter. They reveal clearly the extent of our ignorance. The nature of none of the many intermediates and enzymes which must play a part in photosynthesis has as yet been established, with the sole exception of that of chlorophyll.

* We must emphasize, however, that this is no support for the old theory assuming some relationship between the normal respiration and photosynthesis in plants.

VIII. SUMMARY

During the last decade it has been recognized that the process of photosynthesis in green plants is unique, not because it involves a complicated photochemical decomposition of carbon dioxide for which there is no analogy in the organic world, but because it combines in a unique manner a number of processes each of which may be found in other living cells. If we turn from green plants to purple bacteria, for instance, we find that radiant energy is utilized for the reduction of carbon dioxide. These organisms, however, cannot use simply water as a hydrogen donor and hence are not able to liberate free oxygen. For the reduction of carbon dioxide they depend, in addition to light, upon energetically valuable hydrogen donors such as free hydrogen, hydrogen sulphide, or organic acids. The over-all energy balances of these photoreductions are, therefore, much less favourable than that of photosynthesis in plants. If we turn to organisms not sensitive to light, we find that carbon dioxide can be reduced in complete darkness by several species of bacteria and even by some animal tissues. In this case the mechanism is a coupled oxido-reduction in which an excess of hydrogen donors, either of inorganic or organic nature, has to be sacrificed to promote the

'chemosynthesis'. It is clear that such dark reactions lead not to a gain but to an over-all loss of chemical energy.

Recent advances in the field of respiration and fermentation have taught us that despite the infinite variety of metabolic reactions in living cells the principles governing them are few. Accordingly, it is conceivable that the different reactions involving a reduction of carbon dioxide have many important traits in common, and that the study of any one of them may lead to a better understanding of the process of photosynthesis.

The present article is a report on the metabolism of certain unicellular chlorophyllous algae (several species of *Scenedesmus*, *Ankistrodesmus*, *Rhaphidium*) that are able to reduce carbon dioxide either in normal photosynthesis with the evolution of oxygen, or in photoreduction with the absorption of an equivalent amount of hydrogen, or in chemosynthesis with the oxyhydrogen reaction as the driving force. The two latter reactions do not occur under normal aerobic conditions. They can be observed only after a few hours' incubation in hydrogen gas. The anaerobic treatment brings into play a hydrogenase which enables the algae to absorb or to release molecular hydrogen. This metabolic change we call adaptation. The adaptation consists in an enzymatic activation or rearrangement of some of the catalytic systems. It can be inhibited by traces of specific poisons like cyanide and hydroxylamine.

Upon illumination in the adapted state, in presence of hydrogen and carbon dioxide, the algae reduce carbon dioxide with twice the volume of hydrogen, exactly akin to some purple bacteria. This we call photoreduction. The results of experiments with flashing light and with specific poisons show that in photosynthesis and photoreduction the truly photochemical reactions are the same and remain unchanged. Hence, the difference appears to originate from the ways by which the oxidized products of the photochemical reaction are eliminated. In photosynthesis they are decomposed with the liberation of oxygen, in photoreduction they are reduced to water by hydrogen donors.

The adapted state of the algae gives way to normal aerobic conditions not only under the influence of air, but also under the influence of higher light intensities. This we call reversion. It seems that reversion occurs whenever some intermediate oxidized products (which we call 'peroxides', because they must be the precursors to molecular oxygen) accumulate faster than they are reduced by the hydrogenase system. In absence of carbon dioxide no intermediate oxidized products (or 'peroxides') are formed and the adapted state is stable even at high light intensities. If not only carbon dioxide but also hydrogen is absent (i.e. in an atmosphere of pure nitrogen), light causes a release of hydrogen from the adapted algae. The reversion by light in presence of carbon dioxide can be prevented by the action of certain substances like hydroxylamine, or *o*-phenanthroline. In the presence of such inhibitors the adapted algae continue to metabolize like purple bacteria even in very strong light. The light saturation rate of carbon dioxide reduction with hydrogen remains, however, far smaller than the corresponding rate of photosynthesis.

Very small amounts of oxygen prevent adaptation, but up to 10 mm. Hg of this gas are tolerated by adapted algae in hydrogen. The reason is that at these low partial pressures the reduction of oxygen to water in the algae proceeds faster than the reversion. Under optimum conditions the formation of water from the elements, the oxyhydrogen reaction, is coupled with a simultaneous reduction of carbon dioxide, so that one-half molecule of carbon dioxide disappears for each molecule of reduced oxygen. The coupling between oxyhydrogen reaction and the reaction of carbon dioxide is such that in absence of carbon dioxide the first reaction remains incomplete; little more than one molecule of hydrogen is absorbed instead of the two necessary to form water. After a partial reduction of the oxygen leading to the 'peroxide' level, the reaction apparently continues in unknown directions with internal hydrogen donors. Again a different result is obtained after adding specific poisons. In presence of carbon dioxide and of poisons, the amount of hydrogen absorbed indicates the straightforward formation of water.

A critical comparison of the results reported in this article leads to the assumption that the dark reduction of carbon dioxide can be represented very satisfactorily by reversing essential parts of a diagram drawn to describe various partial reactions in photosynthesis and photoreduction.

My thanks are due to Drs J. Franck and D. Goddard, for helpful suggestions. I am particularly indebted to Dr E. Rabinowitch for a very thorough criticism of the manuscript.

IX. REFERENCES

AUFDEMGARTEN, H. (1939): *Planta*, **30**, 343.

BRIGGS, G. E. (1941): *Proc. Roy. Soc.* B, **130**, 24.

BURR, G. O. & MYERS, J. (1940): *J. Gen. Physiol.* **24**, 45.

CLAREN, O. B. (1938): *Liebigs Ann.* **535**, 122.

DOUNCE, A. L. & HOWLAND, J. W. (1943): *Science*, **97**, 21.

DUTTON, H. J. & MANNING, W. (1941): *Amer. J. Bot.* **28**, 516.

EMERSON, R. & ARNOLD, W. (1932): *J. Gen. Physiol.* **16**, 191.

EMERSON, R. & LEWIS, CH. M. (1939): *Amer. J. Bot.* **26**, 808. — (1942): **29**, 121.

FARKAS, A., FARKAS, L. & YUDKIN, J. (1934): *Proc. Roy. Soc.* B, **115**, 373.

FRANCK, J., FRENCH, C. S. & PUCK, T. T. (1941): *J. Phys. Chem.* **45**, 1268.

FRANCK, J. & FRENCH, C. S. (1941): *J. Gen. Physiol.* **25**, 309.

FRANCK, J. & HERZFELD, K. F. (1941): *J. Phys. Chem.* **45**, 978.

FRANCK, L. & GAFFRON, H. (1941): *Advances Enzymology*, **1**, 199.

GAFFRON, H. (1935*a*): *Biochem. Z.* **275**, 301. — (1935*b*): **279**, 1. — (1937): *Naturwissenschaften*, **25**, 460, 715. — (1939*a*): *Nature, Lond.*, **143**, 204. — (1939*b*): *Biolog. Zbl.* **59**, 302. — (1940*a*): *Amer. J. Bot.* **27**, 273. — (1940*b*): **27**, 204. — (1940*c*): *Science*, **91**, 529. — (1942*a*): *J. Gen. Physiol.* **26**, 195. — (1942*b*): **26**, 241.

GAFFRON, H. & RUBIN, J. (1942): *J. Gen. Physiol.* **26**, 219.

GAFFRON, H. & WOHL, K. (1936): *Naturwissenschaften*, **24**, 81, 103.

GJESSING, E. C. & SUMNER, J. B. (1942): *Archiv. Biochem.* **1**, 1.

HAAS, E., HARRER, C. & HOGNESS, T. R. (1942): *J. Biol. Chem.* **143**, 341.

HARVEY, E. N. (1940): *Living Light*. Princeton Univ. Press.

HES, J. W. (1938*a*): *Nature, Lond.*, **141**, 647. — (1938*b*): *Ann. fermentat.* **4**, 547.
HILL, R. (1939): *Proc. Roy. Soc.* B, **127**, 192.
HILL, R. & SCARISBRICK (1940): *Nature, Lond.* **146**. 61. — (1940): *Proc. Roy. Soc.* B, **129**, 238.
JOHNSTON, E. S. & MYERS, J. (1943): *Ann. Rev. Biochem.* **12**, 473.
KEILIN, D. & HARTREE, E. F. (1935–6): *Proc. Roy. Soc.* B, **119**, 141.
KEMPNER, W. (1933): *Biochem. Z.* **257**, 41.
KEMPNER, W. & KUBOWITZ (1933): *Biochem. Z.* **265**, 245.
KRAHL, M. E. & CLOWES, C. H. A. (1938): *J. Comp. Cell. Physiol.* **11**, 1, 21. — (1940): *J. Gen. Physiol.* **23**, 423.
LIPMANN, F. (1943): *Ann. Rev. Biochem.* **12**, 5.
MCALISTER, E. D. (1930): *J. Gen. Physiol.* **22**, 613.
MCALISTER, E. D. & MYERS, J. (1939): *Smithsonian Inst. Publ., Misc. Coll.*, **99**, no. 6.
MANNING, W. M. (1938): *J. Phys. Chem.* **42**, 815.
MANNING, W. M., STAUFFER, J. F., DUGGAR, B. M. & DANIELS, F. (1938): *J. Amer. Chem. Soc.* **60**, 266.
NAKAMURA, H. (1937): *Acta Phytochim., Tokyo*, **9**, 189; **10**, 211.
NOACK, K., PIRSON, A. & MICHELS, H. (1939): *Naturwissenschaften*, **27**, 645.
ORNSTEIN, L. S., WASSINK, E. C., REMAN, G. H. & VERMEULEN, D. (1938): *Enzymologia*, **5**, 110.
PRATT, R. & TRELEASE, S. F. (1938): *Amer. J. Bot.* **25**, 133.
RABINOWITCH, E. (1943): *Photosynthesis.* New York: Interscience Publ.
RIEKE, F. F. & GAFFRON, H. (1943): *J. Phys. Chem.*, **47**, 299.
ROELOFSEN, P. A. (1935): Dissertation, Utrecht.
RUBEN, S., KAMEN, M. D. & HASSID, W. Z. (1939): *J. Amer. Chem. Soc.* **61**, 661. — (1940): *Proc. Nat. Acad. Sci., Wash.*, **26**, 418.
RUBIN, J. (1941): Dissertation, Univ. Chicago, Illinois.
RUHLAND, W. (1924): *Jb. wiss. Bot.* **63**, 321.
SMITH, J. H. C. (1940): *Plant Physiol.* **15**, 183.
SPOEHR, H. A. (1926): *Photosynthesis.* New York: Chem. Cat. Co.
STEPHENSON, M. (1939): *Bacterial Metabolism.* New York: Longmans, Green and Co.
SUMNER, J. B. (1941): *Advances Enzymology*, **1**, 163.
SWEDIN, B. & THEORELL, H. (1940): *Nature, Lond.*, **145**, 71.
THEORELL, H. (1940): *Ark. Kemi Min. Geol.* **14** B, no. 20.
VAN NIEL, C. B., RUBEN, S., CARSON, S. F., KAMEN, M. D. & FOSTER, J. W. (1942): *Proc. Nat. Acad. Sci., Wash.*, **28**, 8.
VAN NIEL, C. B. (1941): *Advances Enzymology*, **1**, 263.
WASSINK, E. C. & KATZ, E. (1939): *Enzymologia*, **6**, 145.
WASSINK, E. C., VERMEULEN, D., REMAN, G. H. & KATZ, E. (1938): *Enzymologia*, **5**, 100.
WARBURG, O. (1928): *Katalyt. Wirkungen d. Lebendigen Substanz.* Berlin: J. Springer.
WELLER, S. & FRANCK, J. (1941): *J. Phys. Chem.* **45**, 1359.
WERKMAN, C. H. & WOOD, H. G. (1942): *Advances Enzymology*, **2**, 135.
WIELAND, H. & PISTOR, H. T. (1936): *Liebigs Ann.* **522**, 116. — (1938): **535**, 205.
WILSON, J. B. & WILSON, P. W. (1943): *J. Gen. Physiol* **26**, 277.
WOHL, K. (1940): *New Phytol.* **39**, 33. — (1941): **40**, 34. — (1937): *Z. phys. Chem.* B, **37**, 209.
YAMAGATA, S. & NAKAMURA, H. (1938): *Acta Phytochim., Tokyo*, **10**, 297.

THE ECOTYPE

BY J. W. GREGOR, Scottish Plant Breeding Station, Corstorphine, Edinburgh

(*Received* 10 *May* 1943)

I. INTRODUCTION

In attempting to generalize from intraspecific taxonomic data the student of speciation is immediately faced with the difficult task of assessing what biological significance he is going to accord the various classificatory categories. So varied are the interpretations which can be placed upon any one of them that none 'has at the moment any definite meaning at all because each has so many different meanings' (Wilmott, 1939, p. 38). It was for this very reason that Turesson refrained from applying the term variety to the category which is known as the ecotype. When, twenty years ago, he introduced his new term it was thought by some botanists, rather optimistically it is true, that here at last was a variational category, the biological significance of which would not be in doubt. Since the ecotype is now beginning to appear in theoretical discussions relating to speciation it seems a most appropriate time to examine the field of variation covered by this intraspecific concept.

Turesson's cultural experiments with plants collected from many different habitats had led him to conclude that 'the genotypical response of the species population to a definite habitat' results in the formation of distinctive 'habitat types' or *ecotypes*; the ecotype being defined as 'the product arising as a result of the genotypical response of an ecospecies to a particular habitat' (1922, p. 112). The habitat environment is likened to a sieve which sorts out from among the constituents of a population those genotypes which are best fitted to survive. An excellent illustration of the effects of this sifting process on the composition of heterogeneous populations is given by Sylvén (1937). A population of white clover (*Trifolium repens*) raised from directly imported seed gave a relative yield of green matter of 100 units at Svalöf in Sweden, whereas populations raised from first and second generation seed of the same foreign origin harvested at Svalöf yielded 129 and 137 units respectively. These results are attributed to the elimination from the reproductive

populations of the less hardy genotypes. On the other hand, a foreign sample which had attained a greater degree of homozygosity as a result of controlled inbreeding was incapable of responding genotypically to the Svalöf environment and gave the values 100, 100 and 102. In the agricultural literature there are many such examples of selection by the environment. For instance, in Australia, Donald (1941) cites the occurrence of races of subterranean clover (*Trifolium subterraneum*) which vary greatly in earliness of maturity when grown in the same environment. If 'Dwalganup' and 'Mt Barker', early and mid-season races respectively, are mixed and sown at Adelaide, in the course of a few years the resulting population is composed almost entirely of early genotypes. If, however, the same mixture be sown under the higher rainfall conditions of the Adelaide Hills, the population, in an equally short time, will be almost wholly of the mid-season type. The fact that there is a positive correlation ($r = +0{\cdot}69$) between lateness of maturity and the number of flower racemes produced per plant supplies the explanation of this differential selection under the two environments. At Adelaide conditions permit early plants to seed normally, but at the time the mid-season genotypes start flowering the moisture conditions are unfavourable for seed production. On the other hand, in the Adelaide Hills both early and mid-season genotypes find a congenial seeding environment, and since the latter produce many more flowers than the former they increase at the expense of the early genotypes and ultimately dominate the habitat.

It is a matter of very considerable biological interest to know whether 'ecotypic' differentiation of this kind takes place continuously in the direction of environmental change, or whether such differentiation leads to the establishment of geographically delimited 'natural' populations with characteristics peculiarly their own. We shall see that ecotypic differentiation is depicted by Turesson as more or less discontinuous, the individual ecotypes standing out as 'objective' units. Such ecotypes, therefore, are essentially eco-geographical entities. But since a geographical race may also be legitimately interpreted as an eco-geographical entity, the question might well be asked: What, if any, is the difference between the *ecotype* and the geographical *subspecies*? Before attempting to answer this question, however, we shall have to inquire whether less well-defined habitat types or ecotypes have been combined in these objective ecotypes: the evidence suggests that they have. We are accordingly confronted with the problem of deciding whether every population which shows ecotypic differentiation of any degree whatsoever should be regarded as an ecotype. The situation is further complicated by the fact that some authors, among whom is Turesson himself, recognize different kinds of ecotypes, e.g. climatic, soil, and biotic. This circumstance suggests that ecotypic differentiation can be related to individual environmental factors or groups of factors, in which case ecotypes should represent ranges, or stages of variation on genotypic gradients related to particular environmental trends.

While the distributional pattern of ecotypic differentiation is the subject with which we are principally concerned we cannot usefully discuss the part played by environmental selection without at least a brief reference to the role of environment as a spatial 'isolator' of populations, and to the effects which such isolation may have upon the delimitation of ecotypes. Intraspecific variates seldom have the ecological opportunity to establish themselves continuously over large areas. In nature a widespread population constitutes a system of comparatively small and partially isolated populations which tend to intrabreed more than interbreed. As far as plants are concerned even quite a narrow strip of uncongenial environment separating interfertile colonies is a sufficient barrier to check the interflow of genes. It would appear that, especially in regions of marked ecological diversity where colonial formation is often pronounced, this partial isolation of intraspecific populations tends to increase the local effectiveness and rate of ecological selection. Thus we cannot take for granted that the intensity of ecological selection is alone responsible for the degree of ecotypic differentiation exhibited by habitat populations. As in the case of plants 'barrier effects of unsuitable habitats have been described for ecologically specialized species of nearly all groups of animals' (Mayr, 1942, p. 233).

In plants a striking feature of these partially isolated intrabreeding communities or gamodemes (Gilmour & Gregor, 1939, p. 333) is their individuality which often expresses itself in respect of genetic attributes which cannot be related to any observable ecological gradient. That the same local population individuality also is found in populations of motile animals has been ably demonstrated, for instance by Dice (1937, p. 20; 1939) and Diver (1940, p. 303). Since differences between gamodemes are to be found even in environmentally similar habitats (Crampton, 1916, 1925) there is no reason to believe that differentiation in respect of ecologically controlled attributes should not be accompanied by differentiation in respect of ecologically neutral attributes. In fact the colonial structure of wild populations is the one best calculated to facilitate the fortuitous fixation of chance variations.

II. THE ECOSPECIES

Before proceeding to examine the nature of ecotypes we must first of all assess the interpretation to be placed upon the *ecospecies*, a specific category comprising one, or more than one, ecotype. The term ecospecies was introduced in 1922 (Turesson, 1922, p. 102) to cover 'The Linnean species or genotype compounds as they are realised in nature' (1922*a*, p. 344). It would seem that, at the time Turesson

defined the ecospecies, his intention was to draw a distinction between the Linnean species as ecological units (ecospecies), and the genotypical construction of Linnean species (genospecies), a distinction which is now no longer so obvious.

In 1929, however, a change of considerable importance was made in the meaning of the ecospecies when Turesson amplified his original definition and redefined the ecospecies as 'An amphimict-population the constituents of which in nature produce vital and fertile descendants with each other giving rise to less vital or more or less sterile descendants in nature, however, when crossed with constituents of any other population' (1929, p. 333). From the above it will be noticed that the question of inherent barriers to the free exchange of genes between populations is now incorporated in the definition, and further that apomictic populations are definitely excluded. That barriers to the free flow of genes between populations had not up till then entered into the ecospecific definition may be gathered from the statement that in certain localities hybrids between an *Atriplex* ecospecies (comprising *A. praecox* Hulphens, *A. longipes* Drejer, and *A. hastifolium* Salisb.; Turesson, 1922, p. 103) and other species of the genus 'become so frequent that it is sometimes difficult to find representatives of the above-mentioned western type [? ecotype] except in certain places where other species of the genus, such as *A. Babingtoni* and *A. latifolium* grow but sparsely' (1922, p. 105). He goes on to say: 'Here, then, the typicalness of the type depends on the number of other species present on the same spot. When these latter are present in sufficient number, 'swamping' of the variety [? ecotype] in question is very likely to occur' (1922, p. 105).

Later still (Müntzing, Tedin & Turesson, 1931) it is made even clearer that specialization to habitat is not the fundamental basis of ecospecific delimitation. On p. 6 we find that 'These biotype groups [ecospecies] are natural units characterized by balance between their constitution and the environment. As many of these units are specialized to a certain type of habitat they have been called ecospecies.... *Only* sterility and viability distinguish between ecospecies.' That different ecospecies are incapable of freely interchanging genes is explicit in this quotation and we may also conclude that the ability to exchange genes freely is the basic feature of populations within an ecospecies. However, we are still left in some doubt as to whether the assessment of gene interchange is to be made only where populations meet in nature. It will be remembered that the words 'in nature' were retained in the 1929 definition, a circumstance which implies that it is not permissible to regard spatially isolated populations as belonging to the same ecospecies, even when such populations have been shown experimentally to be potentially capable of freely exchanging genes. But this implication is not supported by Turesson's writings. For instance, in presenting evidence 'of the effect of climate in differentiating by selection the plant species into ecotypes' he mentions under the title *Armeria vulgaris* Willd. a number of types [ecotypes] found in Scandinavia and 'a dwarf type of the species inhabiting the Faeroes' (1930, p. 106). Reference is also made to the zonation of *A. vulgaris* populations into different climatic types in regions as spatially isolated as Norway and Greenland (1922*a*, p. 338). Furthermore, he gives data relative to the ecotypic differentiation of *Geum rivale* L., which he regards as an ecospecies (1929, p. 333), in Russia, Germany, Sweden, Norway and Scotland. Since populations of an ecospecies situated respectively on the European Continent and Scotland (*Geum rivale*) have no opportunity in nature of showing whether or not they are capable of producing fertile and vital descendants, we may safely assume that the *potential* ability of representative samples of isolated populations freely to exchange genes is the criterion of common ecospecific membership.

We may conclude then, that populations which are capable of exchanging genes freely, irrespective of their spatial or phylogenetic relationships, belong to the same ecospecies. Conversely populations, including chromosome races (Turesson, 1938, p. 413), with a limited capacity for potential gene exchange belong to different ecospecies. Ecospecies which are incapable of gene interchange, though not necessarily unable to cross, belong to different *coenospecies*. As we have already seen, the name ecospecies has no precise ecological significance but has been applied simply because 'many of the units are specialized to a certain type of habitat'.

The changes which the ecospecies has undergone since 1922 undoubtedly affect the status of the ecotype. Its meaning has, for instance, been interpreted in different ways by Sinskaja and Clausen, though both claim to have accepted Turesson's definition; and, as it happens, both claims are correct. Sinskaja (1935) regards those species in which the main discriminating character is of an ecological order as ecospecies, that is she has adopted Turesson's original definition. On the other hand, the ecospecies of Clausen, Keck & Hiesey (1939) are ecological units only in so far as a partial genetic isolation is usually accompanied by some degree of ecological or geographical isolation; these authors have obviously accepted a later definition, the one of 1929, 1930*a*, p. 514.

As a general rule ecospecies which belong to the same coenospecies do not share each other's territories. This regional arrangement of ecospecies has been well illustrated by Clausen and his associates. In the *Artemisia vulgaris* species complex three ecospecies, one of which comprises two ecotypes, stretch across Central California. They differ in chromosome number and are found geographically arranged as follows: *A. Suksdorfi* ($n=9$) is strictly maritime and is replaced inland by *A. Douglasiana* ($n=27$) which is in turn replaced eastward by *A. ludoviciana* ($n=18$), an ecospecies of the Great Basin and the

Great Plains (Clausen *et al.* 1940, p. 326). Ecospecific differentiation, however, does not always involve different chromosome numbers. For example, *Potentilla glandulosa* together with its allies 'so far as we know, has been able to encircle the world and to inhabit climates ranging from warm temperate to arctic-alpine without change in chromosome number' (Clausen *et al.* 1940, p. 29). *P. glandulosa* is provisionally regarded as an ecospecies of the coenospecies *P. arguta* which comprises two other ecospecies *arguta* and *fissa*.

The appreciation of the existence of ecospecies has an important bearing on the question of what populations are to receive recognition as ecotypes. Failure in this direction is, in Clausen's opinion, one of the reasons why regional ecotypes have not received the attention they deserve. For example it would be legitimate to group most of the American species of *Aquilegia* under a single ecospecies, the several species representing morphologically well-defined ecotypes of subspecies rank. A similar case is that of the British bladder campions, *Silene Cucubalus* and *S. maritima*, mentioned by Turrill (1938, p. 359). Turrill's data indicate that there is no inherent barrier to free gene exchange between these two species. If then we regard *S. Cucubalus* and *S. maritima* not as different species, but as ecotypes of the same ecospecies, we have an inland and a coastal ecotype of the subspecies type visualized by Clausen.

Turesson's definition of the ecospecies agrees very closely with the species definition given by Mayr (1942, p. 120): 'Species are groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups.' This definition of Mayr's also bears a strong resemblance to one given by Dobzhansky in which a species is defined 'as that stage of evolutionary process, at which the once actually or potentially interbreeding array of forms becomes segregated in two or more separate arrays which are physiologically incapable of interbreeding' (see Dobzhansky, 1941, p. 373). It will be noted that while Mayr lays emphasis on the ability of intraspecific groups to interbreed, Dobzhansky on the other hand stresses the incapability of different species to interbreed. Dobzhansky's species are obviously the equivalents of coenospecies, but Mayr's species are closely equivalent to the ecospecies, and only to those coenospecies which happen to possess no ecospecies. Therefore one might justifiably ask: Why not drop the eco-prefix and call the ecospecific units species as Mayr has done? There would certainly seem to be no biologically significant reason why the prefix should be retained except that it does indicate a species which has been taxonomically delimited as a result of experimentation, and for this, if for no other reason, its retention in the meantime would seem to be advisable.

III. THE ECOTYPE

1. *Differentiation in response to climate*

Climate tends to vary more or less continuously, yet even in the case of population differentiation accompanying climatic change Turesson's ecotypes usually stand out as comparatively well-defined objective units.

It would seem that the alpine and coastal populations of *Leontodon autumnalis* in Norway are thus clearly delimited because Turesson is able to say that where the two populations contact each other the alpine forms 'mix' with the coastal forms. Turesson's attitude towards ecotypes differentiated in response to climatic environments is perhaps more clearly revealed in the following quotation: '...owing to differences in climate the biotype compound of the ecospecies is split into smaller groups each specialized for a special type of climate, and these groups remain isolated, since the non-specialized biotypes occasionally appearing in the respective groups are steadily eliminated' (1930, p. 516). Nevertheless, he goes on to say that 'an intermediate biotype population is often found in regions where the transition-area between two extreme types of climate attains a certain extent'. Thus he pictures climatic ecotypes as being zonally arranged along major climatic belts with less clearly defined populations occupying the intervening regions.

The differentiation of altitudinal climatic ecotypes presumably conforms to this general pattern. For instance the lowland population of *Solidago virgaurea* changes its genotypic constitution to become first a distinctive subalpine ecotype and then an alpine ecotype as the species follows an altitudinal climatic gradient upwards to the alpine regions; but whether these distinctions are attained gradually, or by two more or less abrupt steps, unfortunately is not stated. We may take it, however, that Turesson accepts differentiation in response to climate as a regional phenomenon involving the more or less sudden replacement of one comparatively uniform population after another at intervals along the climatic gradient. It is of course not improbable that the ecotypic distinctness has been accentuated by the distance between the points on a climatic gradient at which samples were taken: the greater the distance the greater the ecological difference and consequently the more pronounced the degree of ecotypic differentiation. Little wonder, therefore, that a sample of the southern German ecotype of *Solidago virgaurea* described by Turesson is relatively sharply distinguishable from the ecotype inhabiting southern Sweden. If we take Turesson's evaluation of the effects of climatic control at its face value we arrive at the conclusion that the selective influence of the environment is not sufficiently intense to occasion a genotypic response to minor changes in climate. If this is really true, then only at widely separated points on a climatic gradient, where the necessary degree of climatic divergence is reached to initiate

selection, could differentiation take place. These relatively widely separated localities should accordingly represent the centres from which the climatic ecotypes would acquire their characteristics and from which these characters would spread to meet each other somewhere in the intervening areas.

Turesson stresses the advisability of examining the genotypical constitution of species 'before conclusions are made as to the role played by outer factors in the distribution of a species'. That his studies of altitudinal ecotypic differentiation are of very special interest to phytogeographers there can be no doubt, as they offer a reasonable explanation of the causes for differences in distribution of certain alpine populations. For instance, Turesson collected samples of three species which grow both in Europe and in the East Altaian mountains and he found that an alpine ecotype of all three species occurs in the Altaian mountains, differing in height and earliness from the lowland population of the same region, while in the Alps and in the Carpathians apparently no such ecotype exists. In the Altai the alpine ecotype is remarkably tall, a point suggesting that it has been locally differentiated from the giant types of these species inhabiting the Altaian lowlands. The absence of an alpine ecotype in the Alps and Carpathians is, in Turesson's opinion, due to the absence in these regions of the biotypes necessary for the differentiation of that type (1931, p. 345).

Turesson's observations regarding the nature of climatic ecotypes have been amply confirmed by the work of Clausen, Keck and Hiesey. To take one example, *Potentilla glanulosa* is an ecospecies which has been able to occupy almost the entire transect across central California by a simple differentiation into ecotypes. In the Coast Range is found a distinct climatic ecotype, 'subspecies *typica*'. Another recognizable climatic ecotype, 'subspecies *reflexa*', spatially isolated from *typica*, occurs in the Sierra Nevada foothills, and is overlapped by a mid-altitude meadow ecotype, 'subspecies *Hanseni*'. These two latter ecotypes are separated ecologically, for the meadow race never grows on dry hillsides and the foothills ecotype rarely invades the moister meadows. At a greater elevation two other ecotypes occur, a subalpine and an alpine. 'These form an intergrading series morphologically so that it is impracticable to separate them taxonomically. Together they comprise subspecies *nevadensis*' (Clausen *et al.* 1941, pp. 238 et seq.). In the above examples the areal dispositions of three ecotypes are comparatively well defined, but no line of demarcation separates the fourth and fifth. These two latter ecotypes provide what seems to be an example of ecotypic differentiation unaided by spatial isolation. It is however possible even in such cases, according to Clausen *et al.* (1939, p. 105), to distinguish ecotypes because 'at a distance from the point of contact they are quite pure'. We find, therefore, that just as in the case of Turesson's climatic ecotypes, each of Clausen's climatic ecotypes occupies a coherent climatic zone and contacts neighbouring ecotypes only along the borders of this zone, the centre of the zone acting as a reservoir of relatively uncontaminated gene supply. Dr Clausen informs me that one of the reasons why so many investigators have been unable to recognize ecotypes as regional climatic races is that they investigate too small and too local units.

2. *Differentiation in response to edaphic and biotic environment*

An interesting example of local ecotype formation in response to soil conditions is recorded by Turesson from a much edaphically diversified area of limited extent (35 km. in length) situated in the south-eastern corner of Sweden. Here the arenaceous fields at Vitemölle, a place at the northern end of the area, are occupied by a prostrate dune type of *Hieracium umbellatum*. 'This type is succeeded by the Stenshuvud sea-cliff type south of Kivik, which again is replaced by the first-mentioned type a little further to the south where the sandfields appear. The type in question now flourishes as far south as the extent of the sandfield area, but is replaced by the type of the shifting dunes as soon as the drift sand of the Sandhammar region begins' (1922*a*, p. 337). That these population differences are hereditary and not simply modificatory has been convincingly shown. It is therefore of considerable interest to note that with reference to these localized ecotypes Turesson once again reiterates his conviction 'that in a strictly intermediate area between two habitats of different nature and populated by two different types no intermediate population is found but on the contrary individuals of the two definite types and bastards between them appear' (1922*a*, p. 331). Now at a place called Kylsgård, within the area under discussion, there is found separating the arenaceous fields from the drift sand habitat 'a neutral zone, into which habitats of different nature insensibly grade'. It is therefore important to examine the data relating to this neutral zone, and to compare the characteristics of its occupants with representative ecotypic samples from the arenaceous fields at Vitemölle and the drift sand region at Löderup. The following are the data converted into percentages (1922*a*, p. 332): Vitemölle, 0% erect, 6% ascending and 94% prostrate; the neutral zone at Kylsgård, 9% erect, 36% ascending and 55% prostrate; Löderup, 39% erect, 54% ascending and 6% prostrate. These figures, while bearing out Turesson's contention of a mixture of types, nevertheless indicate that as far as the proportions of these habitus types are concerned the Kylsgård population is also intermediate between his two ecotypes. This example hardly supports the statement that 'where the climate is edaphically diversified several distinct habitat-groups may be found, each marked off by definite and hereditary biotype-characteristics' (1930, p. 516). Thus when Turesson refers to the definiteness of ecotypes we

are entitled to conclude that sometimes at least this definiteness is either a reflexion of the sampling technique or is exaggerated at the expense of transitional populations. Nevertheless where the dividing line between markedly different habitats is sharp, a correspondingly abrupt separation of genotypically different populations may undoubtedly be found. But it in no way follows that because populations occupying distinctive adjacent habitats are relatively abruptly demarcated, that all ecotypes are equally clearly differentiated. Degrees of ecotypic discontinuity and divergence should not be assessed only on the evidence obtained from comparison of extreme habitat populations, but by an examination of differences exhibited by populations inhabiting the various habitats lying between the extremes. On Turesson's own showing, intermediate populations may be found where the area between extreme habitats 'attains a certain extent'. Perhaps the Kylsgård transitional zone is just such an area.

If now we take an example from the north-west coast of the same province of Scania we find that edaphic ecotype differentiation in *Hieracium umbellatum* has also taken place there. In this case samples were taken from three rocky and three sand dune localities which were situated as follows: *rock*, (*a*) Kullen, (*b*) Hofs Haller to the north of Kullen, and (*c*) the west side of Hallands Väderö, a small island lying off the coast about 5 km. from Hofs Haller; *sand dune*, (*a*) Torekov, about 4 km. to the south of Hofs Haller, (*b*) Skelderviken, and (*c*) Nyhamn, about 10 km. south of Kullen. In culture the populations from these two distinctive habitats exhibited genetic differences related to their respective habitats in respect of growth habit and width of leaf. These differences become very apparent when we average the variate data given by Turesson for each 'habitat type'. The average values are: *angle of stem from vertical*, rock samples 40°, dune samples 18°; *leaf width*, rock samples 19 mm., dune samples 8 mm. Incidentally it is of considerable interest to find that while the shape of the inflorescence varied much both within and between five of the six samples, only one shape was found in the material from the sixth, the sample from the isolated island of Hallands Väderö.

Unfortunately, not all edaphic habitats are so clearly defined as the ones we have just been considering, nor for that matter are the ecotypes so readily recognized. The sea plantains of the ecospecies *Plantago maritima* which inhabit the coastal area of Britain provide an instance of more or less continuous ecotypic differentiation (Gregor, 1939). In this case the habitats concerned range from waterlogged mud to well-drained mud. Although sea plantains occur all round the coast of Britain, their actual distribution is essentially colonial and many individual populations belonging to the 'mud' series of habitats stand out as quantitatively differentiated and comparatively well-defined objective units, not only in respect of the edaphically significant character, which is in this case growth habit, but in respect of other characters as well. The growth-habit differences between cultivated samples collected from communities occupying adjacent habitats is sometimes most noticeable, particularly is this so when the habitats concerned contrast strongly and some degree of spatial isolation separates their communities. But such breaks as these in the growth-habit continuum do not indicate breaks between adjacent ecotypes on the edaphic series, for when the habitats are subjectively arranged in proper ecological sequence their occupants show a parallel genotypic gradient involving more or less continuous quantitative changes in growth habit. In other words we find that what in terms of geographical distribution appeared as a confused and to some extent objectively analysable differentiation pattern can be resolved into a presentable example of continuous ecological replacement. The difference between ecotypic differentiation in the sea plantains and the more clear-cut differentiation described in the previous example is apparently only a matter of degree. Accordingly, it makes one wonder whether some at least of Turesson's edaphic ecotypes are not just prominent reference points between which stretch arrays of less distinct ecotypic populations that by their exclusion from ecotypic status tend to throw into relief the more conspicuously differentiated populations.

It is of interest to note that Clausen has not been able to detect edaphic ecotypes in California. He attributes this to the erratic distribution of edaphic populations, and to the swamping effect occasioned by the ample crossing opportunities afforded them by their frequent close proximity to each other. He nevertheless appreciates that populations are sometimes related to soil types, but such relationships are, he thinks, rarely, if ever, ecotypic; instead they have turned out to be ecospecific, e.g. *Viola tricolor-arvensis*, *Dactylis Aschersoniana-glomerata*, and *Phleum nodosum-pratense* are edaphically contrasting ecospecific pairs.

Just as in the case of edaphic ecotypes, differentiation in response to the biotic environment is likely, as a rule, to conform to the geographically 'scattered' distributional pattern. To illustrate the selective control of biotic factors we can hardly do better than take an example from the work of Sinskaja on the differentiation of phytosocial ecotypes or, as she calls them, *synecotypes*. The formation of these synecotypes is visualized by her as proceeding along three main lines: (1) mimicry, (2) differentiation, where equilibrium in a plant association is attained by selection of different forms which utilize space in a different way, and (3) the selection of forms adapted to the environment created by the plant association. Now in Asia Minor there occurs a climatic ecotype of black mustard (*Brassica nigra*) which is spread exclusively in pure crops. However, there is also present in this region, under the same climatic conditions, another quite different type of black mustard.

This second form is never found as a pure crop but only occurs mixed with *Brassica campestris*. Here then we have a phytosocial ecotype selectively adapted to live in association with *B. campestris*, and a phytosocial environment sufficiently strong to overcome the effect of climatic factors and 'to direct selection along quite an opposite line' (Sinskaja, 1931, p. 66).

IV. THE CLASSIFICATION OF ECOTYPIC DIFFERENTIATION

Turesson has recognized four regional ecotypes in Europe, the lowland, the maritime, the subalpine and the alpine, which represent the hereditary adjustments of ecospecies to major regional environments. He has given technical names to three of these ecotypes (1925, p. 222). Oecotypus *salinus* comprises the hereditary habitat types of non-halophytic inland plant species arising in response to saline habitats, while the names *subalpinus* and *alpinus* have been applied respectively to the hereditary habitat types within lowland species differentiated in response to subalpine, and alpine, habitat factors. In addition to these regionally distributed ecotypes he has named two others which must, in view of the nature of their habitats, almost inevitably exhibit a more localized distribution pattern. The hereditary habitat type differentiated within species in genotypical response to the habitat factors of the 'Alfvar' (calcareous rock) and ecologically similar formations is designated oecotypus *campestris*, while those habitat types developed in response to the shifting dune habitat belong to oecotypus *arenarius*.

Notwithstanding what may be called, for want of a more descriptive title, this 'habitat subspecies' conception of ecotypes, Turesson, throughout his writings, is repeatedly drawing distinctions between climatic and edaphic ecotypes, ecotypes which, with some exceptions, e.g. two Siberian ecotypes of the grass *Dactylis glomerata* (1929*a*), are apparently assigned to the various major groupings. For instance, the breadth of his conception of the *salinus* ecotype has necessitated, in the case of the sea pink, *Armeria vulgaris*, the recognition of climatic ecotypes within its borders (1930, p. 105). Similarly two sea-cliff populations, distinct products of genotypical response, which occur respectively on the east and west coasts of Sweden belong to the *salinus* ecotype of *Hieracium umbellatum*. Moreover, that populations belonging to very different 'particular habitats' are embraced by this ecotype is shown by the fact that a cliff population of *Melandrium rubrum* and one of *Lythrum salicaria* from brackish pools are recorded as belonging to the oecotypus *salinus* of their respective species. Thus while Turesson's definition of the ecotype is applicable to any one of the numerous habitat populations which he has described, it is apparent that he prefers to draw attention to a limited number of primary groupings, presumably those populations which he considers show the greatest degree of objective group distribution. We also see that the named ecotypes have reference *either* to some essentially edaphic environment, e.g. *campestris* and *arenarius*, or to some predominantly climatic environment, e.g. *subalpinus* and *alpinus*, or to a combination of both, e.g. *salinus*.

The regional pattern of ecotypic differentiation has also been stressed by Clausen *et al.* On the other hand Clausen's ecotypes are all of one kind, the ecotypes which replace each other regionally in the direction of a progressive change in climatic environment. The characteristics which distinguish these climatic ecotypes are largely physiological as the following example shows: The ecotypes of the Californian Coast Range flower earlier at Stanford (30 metres elevation) than do the mid-altitude ecotypes, but are later when grown at the Mountain Stations (1400 and 3050 m.); their herbage is relatively frost resistant in contrast to the herbage of mid-latitude ecotypes which enter a dormant period when frost begins; at Timberline (3050 m.) Coast Range ecotypes are unable to ripen seed and rapidly decline and die, mid-altitude ecotypes, however, survive at Timberline but either fail to flower or flower too late to produce ripe seed.

The number of climatic ecotypes into which any ecospecies is divided depends largely on the extent of the ecospecific range, and on whether or not its distribution area is much climatically diversified. As a rule the number is small. For instance it only requires six ecotypes of *Potentilla glandulosa* to occupy the climatically diverse region across central California. This relatively small number is attributed to the fact that each ecotype has its 'latitude of tolerance'. Despite the occurrence of intergrading where they meet there would seem little doubt as to the 'objectiveness' of these climatic ecotypes.

As we have seen, not all the Californian ecotypes are accorded taxonomic ranking. Because the subalpine and alpine ecotypes of *Potentilla glandulosa* form an intergrading series morphologically they have been included within the morphologically distinct subspecies *nevadensis*. Nevertheless these two ecotypes are distinguishable and climatically significant units, and as Clausen *et al.* (1941, p. 241) say 'the physiological differentiation within the subspecies, separating an early-flowering, dwarfish and alpine race from a later-flowering, taller, subalpine one, may be just as important as the differences distinguishing this subspecies from the others'.

1. *Ecotypes and subspecies*

What then *is* the difference between ecotypes and subspecies? In the first place we must define what we mean by subspecies. According to Huxley (1939*a*, p. 105) a subspecies is 'a natural or "real" taxonomic unit, in the sense that it is a self-reproducing group with a characteristic geographic distribution, distinguishable from other similar groups by measurable character-differences which can be

determined on any reasonably-sized series. Where genetic analysis is possible, it is interfertile with adjacent subspecies of the same polytypic species'. It will be noticed that the kind of characters which distinguish subspecies are not specified. On the other hand, the definition of a geographical race (subspecies) given by Rensch (1934), though fundamentally the same as Huxley's, quite definitely restricts the use of the term to morphologically differentiated populations, for he says: 'A geographical race is a complex of interbreeding and completely fertile individuals which are morphologically identical or vary only within the limits of individual, ecological and seasonal variability. The typical characters of this group of individuals are genetically fixed and no other geographical race of the same species occurs within the same range' (see Mayr, 1942, p. 106). Dice has approached the problem of subspeciation from a more pronouncedly ecological angle. In his opinion 'the subspecies is best considered to represent an ecological response, largely of an hereditary nature, expressed by those members of a subspecies, which live in a restricted but not always continuous geographic area having a certain type of environment' (1940*a*, p. 292). But while primarily an ecologic unit the subspecies 'must of necessity be based on morphological characters' (1940, p. 220).

Clausen *et al.* adhering to a morphological conception of subspecies answer our question in the following manner. 'The ecotypes within *Potentilla glandulosa* have been worked out experimentally... by determining the number of groups within the species that differ from one another ecologically and physiologically. In the majority of cases these also differ morphologically....The ecotypes have been the basis of our taxonomic treatment, for we consider a morphologically distinguishable ecotype the equivalent of a subspecies. As shown in our treatment of *P. glandulosa* the ecotype (a genetic-ecologic unit) may not always coincide with the subspecies (a systematic unit). This is because ecotypes are not always morphologically distinct, and they must then be retained within one systematic unit or subspecies' (1940, p. 32).

A striking feature of regional population differentiation in California is that ecotypes apparently never overlap the morphological subspecific borders. It would seem, therefore, that the circumstances which occasioned the limits of morphological 'systematic' subspecies have also been responsible for delimiting the boundaries of the physiologically differentiated ecotypes. Of course, that is not to say that the morphologic boundaries are ecotypically determined. For instance in *Sesamum indicum* Sinskaja has shown that tetrameral florets can be distinguished from the ecotypic characters as an historical attribute of wide distribution. Unfortunately, for the systematist these subspecies are not always monotypic, for within a morphological boundary there may occur two or more physiologically differentiated ecotypes which share the subspecific identification marks and may not be themselves distinguishable morphologically.

The ecotype as defined by Turesson is any assemblage of organisms which genotypically reflects the selective action of environment, and may represent anything from a small colonial community to a large regional race. The term ecotype is therefore a general one covering assemblages of very different taxonomic significance. For purposes of theoretical discussion it is, however, highly advisable to consider as 'ecotypic subspecies' those ecotypically differentiated populations which can be shown to possess the fundamental properties of truly objective races, i.e. ecotypes that represent special stages of equilibrium in the differentiation process—'namely that of partial discontinuity' (Huxley, 1939*a*, p. 107). We have seen how Clausen's 'taxonomic' subspecific boundaries and those of some ecotypes coincide. But only when there is this coincidence do the ecotypes happen to be recorded as subspecies, and, what is more, there is nothing to indicate that these 'ecotypic subspecies' are in any way biologically different from orthodox geographical subspecies which have been delimited on morphological criteria alone. This matter will be again referred to at the end of the following section.

2. *Ecotypes and ecoclines*

In regions of considerable ecological diversity subspecies are likely to possess a high variability. 'Such variable subspecies are of course difficult to classify, especially when several opposing ecological tendencies are involved' (Dice, 1940, p. 220). Now Dice has shown that within the Nebraska population of the mouse *Peromyscus maniculatus* a conspicuous trend in respect of pelage colour follows a soil colour gradient. The inhabitants of pale-coloured sand areas are pale in colour, and wherever the soil is darker, the mice tend to be darker also. Crossing this ecologically significant pelage-colour gradient is a wholly independent size gradient for which no ecological relationship has so far been established. 'Indeed the size gradient continues across the sandy hills, just as though they were not there' (1941, p. 15). Differences involving a combination of both colour and size have been used in delimiting three subspecies in Nebraska and the adjoining territories. One of these subspecies, the subspecies of the Nebraska sand hills, also embraces a population from the Bad Lands of South Dakota, a population which has been included on the basis of size characters, despite the fact that it inhabits an area of much variability in soil colour. Ecologically speaking the Nebraska sand hills race is a subspecies representing the 'pale' stage on the ecologically significant pelage colour trend. But it must be remembered that in a given subspecific area only the trend which figures in the subspecies diagnosis is going to receive taxonomic recognition. Other coexistent trends may well remain unrecorded unless they extend beyond the subspecific boundary and happen to be used in the

taxonomic delimitation of other subspecies. That such trends are of common occurrence and may be entirely independent of each other is shown by the vast mass of data relating in particular to bird, mammal, and insect subspecies (see Huxley, 1939, 1942; Mayr, 1942).

The delimitation of ecotypes by Clausen *et al.* has not been complicated by the presence of coexistent ecotypic trends for they have concerned themselves entirely with one variational trend, the trend which follows variation in climate. These authors do, however, mention that besides considerable variation within their ecotypes 'race complexes are also in evidence'. Thus even in California the climatic ecotype apparently does not exhaust the field of ecotypic differentiation. Examples have already been given of ecotypic differentiation in response to edaphic and biotic environments; some of the ecotypes thus formed possess the objective reality of subspecies. However, sometimes, as in the case of the sea plantains, the absence of ecological breaks in a habitat series taken as a whole, and the geographical intricacy of the habitat pattern make the identification of objective ecotypes a practical impossibility. That is not to say that populations following, let us say, a particular edaphic gradient never possess objective reality, for where relatively extreme habitats of the series adjoin, adjacent populations may appear as well differentiated units. For example, two coastal sea plantain populations occupying respectively waterlogged mud and drained 'salting' habitats, separated only by a stream, gave the following average plant height values when seed samples were measured in culture: waterlogged mud, $7{\cdot}1 \pm 0{\cdot}23$ in.; 'salting', $14{\cdot}3 \pm 0{\cdot}28$ in. Nevertheless, when the total array of populations covering this soil gradient was considered as a whole, i.e. in ecological sequence from waterlogged mud to fertile meadows, plant height was found to vary continuously up to $22{\cdot}4 \pm 0{\cdot}42$ in. Obviously in such cases we cannot record ecotypic differentiation in terms of subspecies or micro-subspecies, the distribution pattern of the habitat populations is too complex and the difference between populations adjacent in the ecological sense, too indefinite for that.

Now, if we are to record this kind of ecotypic differentiation we will have to think of differentiation, not in terms of the ecotypic characteristics of geographically adjacent populations, but in terms of the way it is distributed along any particular environmental gradient. Instead therefore of regarding the ecotype objectively as a subspecies we will now have to imagine the ecotype subjectively as a certain range of variation on a genotypic gradient selectively developed in response to a climatic, edaphic, biotic, or for that matter any environmental gradient we may choose to examine.

If we employ the term *ecocline* (Huxley, 1938, p. 219, and 1939) to denote a series of habitat populations showing genotypical gradation related to a particular environmental gradient, and the term ecotype to denote those populations which occupy a particular range of ecotypic variation on an ecocline, it should be possible to record in a summarized and readily understandable manner the trends of ecotypic differentiation. In the same way as Huxley has used prefixes to denote different kinds of clines, so prefixes could be employed to denote different categories of ecoclines and thus permit the independent listing of different ecotypic gradients. The ecoclines and the intraecoclinal ecotypes could be given self-explanatory vernacular names having reference either to the attributes concerned with their recognition, or to the nature of the environmental gradients which they follow. For instance, in the sea plantains the trend of ecotypic differentiation which follows the 'mud' environmental gradient is a trend in growth-habit expression where, on the average, the growth habit of populations becomes more erect and vigorous in the direction of increasing soil fertility. This trend has been called the growth-habit ecocline, and it has been divided for reference purposes into three ecotypically significant ranges, the predominantly decumbent, ascending, and erect ecotypes.

The ecoclines represent ecological subcategories of ecospecies, and the reference populations within them we might conveniently call *ecoclinal ecotypes*. A habitat population which belongs to a particular ecoclinal ecotype by virtue of possessing certain characteristics may at the same time share its ecotypic status with ecotypes on other ecoclines by virtue of carrying their characteristics in addition. Thus the number of times a given areal population can represent a range, or stage, on an ecocline depends on the number of independent and ecologically significant genotypic trends found crossing its territory.

It should perhaps be emphasized that the authenticity of a population's ecological significance can be convincingly shown only when it is possible to refer it to a range, or stage, on an ecocline. This observation applies with equal force to subjectively established ecotypes of the kind we have just been discussing, and to objective ecotypes of the kind recorded by Clausen *et al.* For instance unless these authors had been able to relate the trend of genotypic change to the climatic gradient the existence of the Californian populations as ecotypes could never have been definitely established. But as we have already decided that the Californian populations are truly objective natural populations of subspecific standing they appear on the climatic ecocline, not as subjectively delimited ecotypic ranges, but as objective stages or subspecies. We are therefore faced with the following alternatives: either we should apply the term subspecies to 'objective' ecological units and reserve the term ecotype for the 'subjective' categories, or we should restrict the term ecotype to the former and find a new term to cover the latter categories. The first alternative would seem to be most sensible, and the one that ought to be adopted.

But just as the component populations of the eco-species cannot be determined by inspection alone, so the ecotypic or *ecoclinal subspecies* requires an experimental technique for its delimitation as a taxonomic unit. Therefore, for technical reasons and not least for purposes of discussion, the identity of the ecoclinal subspecies should not be taxonomically merged with that of the *geographical subspecies*. Perhaps, as suggested by Huxley (1942, p. 406), some such method as the prefixing to the subspecific name of the letters 'E', and 'G', would serve to record the distinction.

The *ecoclinal ecotype*, on the other hand, should be regarded as a subjective category of a subsidiary and complementary scheme of classification designed to record the general trends of ecotypic differentiation, and as such it should not receive a technical name. If then an ecotype is to represent a range of ecotypic variation on an ecocline the use of the term in its wider sense is to be deprecated. Where it is found necessary to refer loosely to ecotypically differentiated populations some such non-committal term as ecodeme should be used, thereby avoiding in the future the obvious confusion which has arisen by the indiscriminate employment of the term ecotype in the past.

V. SUMMARY

1. Ecotypic differentiation is hereditary differentiation in respect of morphological and/or physiological attributes occasioned by the selective action of the habitat environment.

2. Ecotypic differentiation in response to one set of environmental conditions is liable to proceed independently of that in response to another set of conditions. Regularities in ecotypic differentiation are therefore best studied in relation to particular environmental trends.

3. Ecotypic differentiation may be regional or strictly local.

4. Regional ecotypic differentiation occurs most frequently in species with wide latitudinal and/or altitudinal distributions, i.e. it appears as a response to variations in climate.

5. A feature of ecotypic differentiation in response to climate is that the range of variation is not usually continuous in the direction of a gradual change in climate, but instead a limited number of populations is formed each with its own climatic 'latitude of tolerance'.

6. Where neighbouring climatic populations meet they lose their individuality, as a rule, only along comparatively narrow zones of intergradation. The limits of the areas of intergradation cannot, however, always be assessed with accuracy as the effects of hybridization 'diminish often at some distance from the point of contact' (Clausen *et al.*), and some genes may travel far.

7. Local ecotypic differentiation may also be characterized by the formation of relatively uniform populations separated by intergradation zones of greater variability. But when adjacent populations occupy small and contrasting habitats the evidence suggests that the effects of hybridization may extend over the entire communities without completely obliterating their respective identities.

8. The local establishment of ecotypically individualistic communities is considerably facilitated where a colonial population structure is pronounced.

9. Ecotypically differentiated populations may coincide with geographically differentiated *subspecies*. When they do coincide with subspecies, or when they do not coincide but still exhibit the fundamental features of truly objective populations, i.e. have definable borders and some degree of ecotypic individuality, they ought to be taxonomically recorded at the subspecies level. The use of the term 'ecoclinal subspecies' has been suggested as a convenient way of making the necessary distinction between experimentally delimited ecological subspecies and orthodox subspecies.

10. Ecoclinal subspecies should receive taxonomic names.

11. Ecotypic differentiation related to particular environmental trends but which can only be recorded in terms of subjectively delimited ranges of ecotypic variation, is not 'subspecific', and populations assigned to a subjective range should not be accorded subspecific status. Such populations have been referred to as 'ecoclinal ecotypes'.

12. Ecoclinal ecotypes should receive only vernacular titles.

13. The general term ecotype as defined by Turesson is applicable to any 'product' of ecotypic differentiation.

14. When it is necessary to refer to habitat populations the ecotypic significance of which has not been established the term ecotype should be avoided, and some non-committal term such as ecodeme should be used instead.

The writer wishes to take this opportunity of thanking Dr Julian Huxley and Dr Irene Manton for suggestions which have been most helpful in deciding the scope and nature of the present review.

VI. REFERENCES

Clausen, J., Keck, D. D. & Hiesey, W. M. (1939): *Amer. J. Bot.* **26**, 103–6. — (1940): *Publ. Carneg. Instn*, no. 520. — (1941): *Amer. Nat.* **75**, 231–50.

Crampton, H. E. (1916): *Publ. Carneg. Instn*, no. 228. — (1925): no. 228A.

Dice, L. R. (1937): *Occ. Pap. Mus. Zool. Univ. Mich.* no. 352. — (1939): *Contr. Lab. Vertebr. Genet. Univ. Mich.* no. 12. — (1940): *Amer. Nat.* **74**, 212–21. — (1940*a*): **74**, 289–98. — (1941): *Contr. Lab. Vertebr. Genet. Univ. Mich.* no. 15.

Diver, C. (1940): *The New Systematics*, pp. 303–28. Edited by J. S. Huxley. Oxford Univ. Press.

Dobzhansky, T. (1941): *Genetics and the Origin of Species*, 2nd ed. New York: Columbia Univ. Press.

Donald, C. M. (1941): *University of Sydney Publ.* pp. 1–108.

Gilmour, J. S. L. & Gregor, J. W. (1939): *Nature, Lond.*, **144**, 333.

Gregor, J. W. (1939): *New Phytol.* **38**, 293–322.

Huxley, J. S. (1938): *Nature, Lond.*, **142**, 219. — (1939): *Bijdr. Dierk.* **27**, 491–520. — (1939*a*): *Proc. Linn. Soc. Lond.* 151st Session, pp. 105–14. — (1942): *Evolution: The Modern Synthesis*. London: Allen and Unwin Ltd.

MAYR, E. (1942): *Systematics and the Origin of Species.* New York: Columbia Univ. Press.
MÜNTZING, A., TEDIN, O. & TURESSON, G. (1931): *Hereditas, Lund,* **15**, 1–12.
RENSCH, B. (1934): *Kurze Anweisung für zoologisch-systematische Studien.* Leipzig: Akademische Verlagsgesellschaft.
SINSKAJA, E. N. (1931): *Bull. Appl. Bot. Genet. Pl. Breed.* **25**, 1–97. — (1935): *Bull. Appl. Bot. Select.* Supplement 73, 77–124.
SYLVÉN, N. (1937): *Imp. Bur. Pl. Genet. Herbage Plants, Bull* 21.
TURESSON, G. (1922): *Hereditas, Lund,* **3**, 100–13. — (1922*a*): **3**, 211–350. — (1925): **6**, 147–236. — (1929): **12**, 323–34. — (1929*a*): **12**, 335–51. — (1930): **14**, 99–152. — (1930*a*): **14**, 511–18. — (1931): **15**, 329–46. — (1938): *Ann. Agric. Coll. Sweden,* **5**, 405–16.
TURRILL, W. B. (1938): *Biol. Rev.* **13**, 342–73.
WILMOTT, A. J. (1939): *Proc. Linn. Soc. Lond.* 151st Session, pp. 96–9.

THE PRESENT STATUS OF THE NEURONE THEORY

BY JOSÉ F. NONIDEZ, Cornell University Medical College, New York City

(*Received* 30 *June* 1943)

I. INTRODUCTION

At the time of Cajal's posthumous publication (1934) the view that nerve cells are independent units had won widespread acceptance. However, the adherents to the opposite point of view (existence of a neurosyncytium) had not been idle. In the years following the publication of Held's critique of the neurone theory (1929) the results of application of the method of Bielschowsky and its modifications to the peripheral nervous system has gradually led to a revival of the old controversy of neuronism versus reticularism. Finally, in 1937, at the meeting of the Anatomische Gesellschaft in Koenigsberg, was sounded what some reticularists (Boeke, Bethe, Schroeder, Bauer) believed to be the death knell of the neurone theory. That the foundations of neuronism were not even shaken by the evidence presented will be shown in the following pages. Inasmuch as the older facts in favour of the neurone theory are adequately reviewed in Cajal's above-mentioned article and in his chapter 'Die Neuronenlehre' in Bumke & Foerster's *Handbuch der Neurologie* (1935), the present article will deal chiefly with the more recent literature.

II. HISTOGENESIS AND GROWTH *IN VITRO*

The study of the histogenesis of the nervous system has yielded important information on the independence of the neurones from the earliest stages of their differentiation. Recent work in this field has corroborated earlier observations, the papers by Windle (1932, 1933), Angulo (1939), Kimmel (1940), Romanes (1941) and Barron (1943) deserving special mention.

In an article entitled 'Der Neuroncytiumbegriff' Bauer (1938) gives an account of the early development of the nerves which differs from the classical descriptions. Following the ideas of Held, Bauer describes an 'epithelial' connective tissue resulting from the merging of cytoplasmic prolongations of the cells of the ectoderm, the neural tube, the endoderm and the mesoderm into a meshwork. The axonic processes of the ventral (motor) neuroblasts enter primary nervous pathways (neurodesmes) of the non-nucleated 'epithelial' connective tissue meshwork, and growing within them they reach their destination in the innervated organs. The development of the nerve fibres is thus an intraprotoplasmic process. (The two figures of embryos —chick and amphibian—showing the 'epithelial' connective tissue are, significantly enough, diagrammatic and do not include the neural crest.) Later on, according to Bauer, mesenchyme cells move into the 'epithelial' connective tissue. The growth of the axonic prolongations of the neuroblasts into the neurodesmes is hypothetical; in early embryos mesenchyme cells are already so numerous that all that can be said is that the growing tips of the axons emerging from the cord find their way among the mesenchymal elements and that they are not ensheathed by them.

The application of the methods of tissue culture to embryonic nerve tissues has received increasing attention as a source of information on the development and structure of the nervous system. The presence of anastomoses in the explanted neuroblasts and between the growing fibres has been emphasized by the supporters of the hypothesis of the neurosyncytium, forgetting that conditions in cultures are vastly different from those that obtain in the embryo. Furthermore, a review of the literature on the subject shows that the presence of anastomoses *in vitro* is by no means universal and that it depends on several factors among which one of the most important is the part of the nervous system under cultivation. According to Levi (1941) anastomoses seldom or never occur in cultures of embryonic cerebrospinal ganglia except among the

finest branches; even so the fused branches sometimes resume their independence, a fact confirmed in moving pictures by Levi & Delorenzi (1934) and by Huzella (1938). Observations by Weiss & Wang (1936) and Szepsenwol (1939) on spinal ganglia of chick embryos show that the explanted cells undergo normal development to the point of becoming monopolar; no mention of anastomoses is found in these articles. Similarly, Meyer & Jablonski (1937) while calling attention to certain changes in the size and shape of the spinal ganglion cells after long periods of cultivation do not report the fusion of their processes.

Esaki (1929) and Levi (1941) have shown that anastomoses never occur in cultures of the fore- and midbrain, and that they are rare in those of the hindbrain and spinal cord in which they are found in young embryos only. Grigorjeff (1931, 1933) has observed fusion of the processes of neuroblasts, but he also states that in cultures of chick-embryo brain the growing fibres do not fuse with each other, and that in areas with complex plexiform arrangement it is very difficult to be sure that anastomoses actually occur. Similar observations have been reported by Lazarenko (1931), who claims that it is possible to differentiate the axon from the dendrites in scattered neuroblasts. Mihálik (1932) reported the presence of nerve processes which have four to eight thin branches and these sometimes wind themselves around a neuroblast in the fashion of a synapse. In a later article (1935) the same author denies the occurrence of permanent anastomoses and concludes that the facts revealed by the study of tissue cultures lend strong support to the neurone theory because the genetic, anatomical and trophic individuality of the neurones can be confirmed in this living material. Serebrjakow (1934) emphasizes that conditions *in vitro* are very different from those in the embryo and, although he illustrates nerve processes fused in a network, he hesitates to accept this as proof of the syncytial nature of the nervous system. In contrast with the preceding observations Bauer (1938, 1940) claims that anastomoses are not only numerous but permanent and that their presence proves the continuity of the neurones in the adult.

In cultures of sympathetic ganglia the growing prolongations of the neuroblasts anastomose more profusely than in explants of cerebrospinal ganglia (Szantroch, 1933*a*, *b*; Levi & Delorenzi, 1934), as reported previously by Lewis & Lewis (1912). This behaviour may have influenced Szantroch (1935) in his interpretation of the structure of the sympathetic ganglia because he claims that their neurones (which he believes to be of mesodermal origin) are arranged in a syncytium from the early stages of their differentiation. This is in contrast with the findings of De Castro (1922, 1932), Juba (1937) and Mihálik (1940), who also studied the development of the sympathetic system with silver techniques and confirmed the existence of independent neuroblasts. That the latter (sympathoblasts) are independent cells is further shown by their migration from the place of origin (the neural crest, or the spinal cord, or both the crest and the cord, according to some authors) to their ultimate destination in the region in which the sympathetic chain will develop.

From the preceding considerations, it is clear that, contrary to the claims of Stöhr (1937), Bauer (1938, 1940) and Boeke (1938*a*, 1940), the evidence derived from studies on tissue culture of the embryonic nervous system does not support the view that growing nerve fibres anastomose freely. In a similar way Boeke's contention (1938, 1940) that growing nerve fibres in the embryo pass through a syncytial intermediate stage before the final modelling of the nerve endings is not backed by the evidence under consideration. In view of these facts Stöhr's statement (1937) that in regard to the capacity of formation of networks by the neuroplasm there is considerable agreement between the results of tissue culture and those obtained with the Bielschowsky technique is rather surprising. Levi (1941), with long experience in the field of tissue culture of the nervous system, remarks that true anastomoses undoubtedly occurring in certain areas of the explant are due in a large measure to the properties of the media employed and that they are of transitory character. He also expresses the view that the data derived from the study of nerve tissue *in vitro* do not favour the neurosyncytial theory.

III. INTERNEURONAL RELATIONSHIPS: THE SYNAPSE

As far as I am aware the only claim advanced against the neurone theory in connexion with the structure of the adult nerve centres since the publication of the papers of Tiegs (1927, 1931) and Held (1929) is based on observations on the human cerebral cortex (Bauer, 1940). Bauer's findings are confirmatory of the observations of Held on the cerebellum and of Akkeringa (1934) on the retina. Accordingly, he believes that the finest twigs of the dendrites and axons merge with a pale meshwork (in Bielschowsky preparations) which connects the neurones with the glia; further, that the neurofibrils course within this meshwork or *Grundnetz* and pass from neurone to neurone. Against this extreme reticularist viewpoint recent work on the structure of the synapse in the nerve centres and the changes it undergoes under experimental and pathological conditions may be mentioned. The papers by Barnard (1940), Barr (1939–40, 1940), Bodian (1937, 1940), Gibson (1937), Glees (1941), Glees & Clark (1941), Hoff (1932*a*, *b*, 1934), Lorente de Nó (1938), Minckler (1940, 1941, 1942), Phalen & Davenport (1937–8), Schadewald (1940), Schimert (1937–8, 1939*b*) and Silver (1942), contain observations which exclude the possibility that the *boutons terminaux* or synaptic knobs are artifacts because they have been seen in material prepared with diverse silver techniques. This is in contrast to the observations of the anti-neuronists,

based almost without exception on the study of slides prepared with the method of Bielschowsky. Even this technique, as will be pointed out later, will impregnate *boutons* not only in the nerve centres but also in the ganglia of the autonomic nervous system. In the light of the observations by the authors mentioned above, the findings of Bauer are far from convincing. Indeed, Polyák's comment (1941) on Akkeringa's 'demonstration' of the retinal glial-neurofibrillar syncytium—'a phantasmagoria fathered by prejudice and mothered by lack of experience'—could be applied to them. The work of Bartelmez & Hoerr (1933) and Bodian (1937, 1940) may be mentioned as among the most satisfactory demonstrations of the discontinuity of the neurofibrils at the synapses in the central nervous system.

In the field of the autonomic nervous system the continuity of the neurones has also been emphasized by the adherents to the neurosyncytial theory. Boeke's observations (1935) on the enteric plexus of *Amphioxus* not only show anastomoses of the processes of the ganglion cells but he claims that there is continuity of the neurofibrils of the preganglionic fibres and the ganglion cells at the synapses. Szantroch (1935), in a study of the development of the sympathetic nervous system finds that the nerve cells are connected with each other by means of their processes. Harting (1930) has also reported anastomoses of the processes of the ganglion cells in the gall bladder. Likewise, Stöhr (1937) has emphasized this feature and he gives illustrations showing the passage of neurofibrils from the dendrites to the cell bodies of the ganglion cells of the plexus of Auerbach. While it is undoubtedly true that examples of continuity of ganglion cells can be found, especially in the enteric plexuses (the so-called interstitial cells), the conclusion that the anastomoses are the rule and that they represent a fundamental arrangement of the nervous system is not in accord with the facts. In intact ganglia it is often difficult to follow the dendrites to their termination because of the elaborate plexus formed by the smaller dendritic branches and the arborizations of the preganglionic fibres. In ganglia in which the preganglionic nerve fibres have been allowed to degenerate, the free termination of the dendrites of well impregnated neurones are easier to perceive (Nonidez & Hare, 1942).

(1) *Demonstration of the synapse.* The existence of a syncytial arrangement of the nervous system implies the absence of synapses; the terminal *boutons* and other synaptic endings described in the literature are regarded by the anti-neuronists as artifacts or as the result of defective impregnation. As mentioned above there are numerous histological observations on the synapse in the vertebrate central nervous system and a good deal of physiological evidence, well summarized in a recent review on the subject by Bodian (1942).

In the case of the autonomic nervous system the evidence on the existence of synaptic endings is more conclusive because in favourable cases it has been obtained in the living animal. In this respect the observations of Fedorow (1935) deserve special mention. This investigator has been able to study the synapse in the ganglia of the living frog's heart and has observed and photographed the terminal *boutons* before and after adding a dilute solution of methylene blue, which stains them as well as the branches of the preganglionic arborization wound around the cell bodies of the neurones. Furthermore, Fedorow has observed the results of various influences (acidosis and alkalosis, hyper- and hypotonia, mechanical irritation, electrical excitation) on the morphology of the *boutons*, thus offering conclusive proof of their reality. The same author, with Matwejewa (1935) has studied the structure of the synapse in the ganglia of the heart and bladder of the frog, using the Bielschowsky-Gros method as well as the methylene blue technique; these two procedures give identical images. That the terminal *boutons* are artifacts due to the inadequacy of our silver impregnation methods, as claimed by Szantroch (1935), seems absurd in the light of the observations quoted.

Histological evidence on the existence of the synapse in the autonomic nervous system has also been presented by De Castro (1922, 1930, 1932, 1936–7), Lawrentjew (1924, 1925, 1929*a*, *b*, 1934), Serebrjakow (1929, 1936), Iwanow & Radostina (1932–3), Dolgo-Saburoff (1936), Nonidez (1937*b*) and Herzog & Günther (1938). It is worth noting that Lawrentjew, Fedorow & Matwejewa, and Herzog & Günther have demonstrated the presence of synaptic endings with the Bielschowsky-Gros technique, the same procedure that in the hands of Stöhr and his school, and of Szantroch (1935), shows 'clearly' the continuity of the neurofibrils of the finer branches of the preganglionic fibres with the neurofibrils of the ganglion cells.

The contention of Stöhr and his students has been that the nerve fibres entering the ganglion, after branching repeatedly give rise to a narrow-meshed network composed of neurofibrils ('terminal' reticulum) which pass into the cell bodies and dendrites of the ganglion cells. The terminal rings, *boutons* and small reticulated enlargements reported by other investigators are due to incomplete impregnation and therefore should be disregarded. Following this line of reasoning it would also be necessary to assume that the *boutons* so convincingly demonstrated by Fedorow (1935) in living ganglia are artifacts, and that their vital staining with methylene blue and their increased coloration upon electrical stimulation of the nerve fibres *in vivo* when neutral red has been added to the preparation are of no significance. (It is interesting to note in this connexion that Bauer (1938) in his article on the neurencytium does not mention Fedorow's observations even though he emphasizes the value of the evidence gathered from studies on nerve tissue

culture. Stöhr (1937), in his section 'Anatomische Bemerkungen zur Frage der Neuronlehre im vegetativen Nervensystem', likewise fails to discuss Fedorow's important observations, although he includes a reference to his work in his literature list). That the network or 'terminal' reticulum of Stöhr and his school is not nervous is shown by the fact that it can be brought about by increasing the concentration of the formalin solution used for fixation (30–40 % instead of the customary 10–15 %) as shown by Lawrentjew (1939, cited by Gibson, 1940). Kolossow & Polykarpowa (1935) have found this reticulum after the preganglionic fibres have been destroyed. Nonidez (1937*b*) has demonstrated a similar structure with silver carbonate, which does not stain nerve cells and fibres.

More recently Stöhr (1939*a*) has given a description of the structure of the sympathetic ganglia which is at variance with all previous observations, including his own and those of Szantroch. According to Stöhr there occurs in these ganglia a plasmodium of cells (*Nebenzellenplasmodium*) which he regards as a tissue *sui generis*. It contains scattered Nissl bodies and, as a capsular plasmodium (*Hüllplasmodium*), it is in close relation with the neuroplasm of the ganglion cells. The plasmodium contains a network of very fine collagen bundles and a meshwork which stains with Rio Hortega's method (silver carbonate) and that Stöhr considers as of probable neuroglial nature. The plasmodium is continuous with the processes of the ganglion cells, thus constituting a morphologically as well as physiologically inseparable unit. This concept of the structure of sympathetic ganglia is not supported by observations on ganglia prepared with the Cajal silver nitrate technique. Furthermore, when as Stöhr has done, two silver impregnation methods are superimposed in the same preparation, the resulting image defies accurate analysis because in addition to the impregnation of the nerve cells and their processes with the Bielschowsky-Gros technique, the Rio-Hortega method stains beautifully the argyrophil connective tissue fibres within the ganglion (Nonidez, 1937*b*).

Boeke (1940), while claiming that certain nerve cells of the sympathetic nervous system are syncytially arranged—a feature previously reported by several authors—accepts the existence of the synapse in a physiological sense. Contrary to Stöhr and his students he does not deny the reality of the synaptic structures that have been described, for he says: 'It is not necessary that these synaptic junctions should have the form of minute rings or net-like endings everywhere such as were described so convincingly by Lawrentjew in 1935. They may be in the form of larger net-like expansions of the neurofibrillar structure or what I have called *wirksame Strecke*, long-drawn-out junctional tissue' (1940, p. 149). ''That a synapse between ganglion cells may be built entirely of living substance with protoplasmic continuity, but with different staining qualities is to be seen in the sympathetic ganglion cells covering the wall of the gut of *Amphioxus lanceolatus*.' Further he states: 'Even in the sympathetic end-formation with its reticular structure a synaptic arrangement remains necessary, but as soon as we recognize that this synaptic arrangement consists of living alterable substance, without intercellular membranous barriers in which it would be impossible to locate the extremely delicate physiological properties we ascribe to synapses, we have to acknowledge that the classical neurone theory has to be abandoned and replaced by a new formula' (1940, p. 155). The conciliatory attitude of Boeke in trying to bridge the gap that separates the neuronists from the reticularists on the subject of the synapse adds little to the solution of the problem. Few neuronists believe at the present time that the synaptic membrane or synaptolemma (Bodian) is an inert substance, a mere *ciment unitif*, but, on the other hand, the assumption that it is a highly organized intermediary formation with specific characteristics does not rest on a histological basis.

To sum up, there is histological evidence on the one hand that the relation of the terminals of the axon to the cell on which it ends is one of mere contact, while on the other hand there are numerous observations favouring the view of continuity of the nerve cells, i.e. passage of neurofibrils from the terminal branches of the axon to the neurone receiving the impulses. The second viewpoint meets a serious setback in the demonstration of the synaptic knobs in the living ganglia and in their staining with dilute solutions of methylene blue, thus removing the objection that the knobs may be artifacts produced by precipitation of metallic salts. It is further weakened by the fact that the same technique used to prove the continuity of the neurofibrils—the method of Bielschowsky-Gros—in the hands of other investigators impregnates clearly the terminals of the axons. The latter can also be demonstrated with various silver techniques either when applied to blocks of nerve tissue or to sections after embedding in paraffin.

(2) *Experimental degeneration of the synaptic endings.* If, as claimed by the neuronists, the terminal *boutons* and other synaptic endings are terminations of branches of axons of other neurones, it follows that they will degenerate after section of these axons. Furthermore, since the neurofibrils of the *boutons* are not continuous with those of the neurone with which they are in synaptic contact, the degeneration should not affect the latter. In very young animals, however, the cell in contact with the nerve endings atrophies or fails to continue its postnatal development; this may be due either to lack of tonic impulses necessary for full differentiation of the young neurone or to disuse atrophy. The latter explanation might also apply to those cases in the adult in which transneuronal effects of axon section are noticeable after several months, as for instance, the atrophy of the neurones of the lateral geniculate body following optic nerve section

(Minkowski, 1920; Balado & Franke, 1930; Clark & Penman, 1934).

Degeneration of the synaptic endings in the central nervous system has been reported by Hoff (1932 *a*, *b*), Hoff & Hoff (1934), Foerster, Gagel & Sheehan (1933), Snider (1936), Gibson (1937), Schimert (1937–8, 1939*b*), Glees (1941), Glees & Clark (1941) and Schadewald (1940) after operative procedures in animals, and by Minckler (1942) in the human cord under pathological conditions. The validity of the observations of some of these investigators is open to question because of difficulties of impregnation, the presence of great numbers of synaptic endings in many of the neurones, and the polymorphism of the *boutons*, which makes difficult the recognition of degenerative changes.

In the case of the autonomic nervous system the evidence is more convincing because the bundles of preganglionic fibres ending in the ganglia can be easily interrupted, and also because the ganglion cells do not receive impulses from so many different sources as the neurones of the nerve centres. The papers of Lawrentjew (1924, 1925, 1929, 1934), De Castro (1930, 1934, 1936–7), Kolossow (1932–3), Kolossow & Sabussow (1932), Kolossow, Sabussow & Iwanow (1932), Baron (1934), Fedorow (1935), Fedorow & Matwejewa (1935, 1936–7), Maksudowa (1936), Kolossow & Polykarpowa (1936), Nonidez (1937*b*), Herzog & Günther (1938) and Gibson (1940) contain observations and figures showing the gradual breakdown and disappearance of the synaptic endings. Against this positive evidence the reticularists have no experimental data to offer. Since Stöhr and his students firmly believe that the preganglionic fibres do not end in contact with the ganglion cells but that they break up into a delicate 'terminal' reticulum the neurofibrils of which are continuous with those of the neurones, one would expect that this important matter would have been investigated. Disappearance of the 'terminal' reticulum and degenerative changes in the neurones after section of the preganglionics should be convincing evidence of the nervous nature of the junctional meshwork. The proponents of the neurosyncytial theory are silent on this point. Harting (1934), however, states that many ganglion cells of the Auerbach plexus of the rabbit's oesophagus show pathological changes after unilateral section of the vagus, a result which disagrees with the reports of other authors who have studied the effects of vagus section, or of section of the sacral parasympathetics, on the ganglia (Lawrentjew, 1929*a*, *b*; Lawrentjew & Naiditsch, 1932–3; Kolossow, 1932–3; Kolossow *et al.* 1932; Fedorow & Matwejewa, 1935, 1936–7; Maksudowa, 1936; Kolossow & Polykarpowa, 1936). Ottaviani (1937), who also describes a preganglionic 'terminal' reticulum in the ganglia of the oesophagus, does not mention the effects of section of the vagus on the ganglion cells.

Physiological proof of the normal condition of the neurones of a ganglion after degeneration of the synaptic endings has been furnished by Gibson (1940). This investigator found that 8 days after section of the cervical sympathetic trunk antidromic impulses, backfired into the cells of the superior cervical ganglion *via* their axons, produced normal action currents from the cells. 'Thus the ganglion cells themselves were not affected by degeneration of the preganglionic fibres' (p. 241).

(3) *Regeneration of synaptic endings.* Since, as stated above, the degenerative changes brought about by section of the preganglionic fibres do not extend beyond the synaptic endings, and since the neurones with which they are in contact remain normal for a period of several weeks or months, regeneration of the preganglionics and their endings should result in a resumption of functional activity of the neurones. Restoration of function was found by Langley (1897)* as early as 24 days after operation, but the effects of stimulation were not as marked as in the intact animal. De Castro (1930) did not notice a return of normal functional activity until 40–50 days after section of the cervical sympathetic, but a partial recuperation was apparent in some cases 15–20 days after section. Hinsey, Phillips & Hare (1939) found functional regeneration of the preganglionic fibres to the ganglia supplying the forelimbs of cats within 36–61 days, and Gibson (1940) working with the superior cervical ganglion elicited responses from its neurones 44 days after operation.

The first histological observations on the regeneration of the intraganglionic branches of the preganglionic fibres are those of Lawrentjew (1925), but in view of later work by De Castro (1930) it is questionable whether the Russian neurohistologist has actually seen the regeneration of the synaptic endings in the superior cervical ganglion since he allowed only a few days (up to 9) for regeneration. De Castro was not able to see regenerated synaptic endings until 12–17 days after operation. In a later paper Lawrentjew (1934) has confirmed (with the Bielschowsky-Gros technique) the findings of De Castro, adding numerous observations to his previous study. Gibson (1940) first saw regeneration of *boutons* 44 days after operation. The figures of the authors mentioned show clearly that the regenerated *boutons* do not differ from those found in intact ganglia.

It is worth noting that regeneration of the synaptic endings also takes place when two different nerves are sutured (heterogenetic regeneration) as for instance the vagus and the cervical sympathetic. The physiological evidence of regeneration in these cases is attested in the work of Langley (1898)* and in more recent contributions by several physiologists (Tsukaguchi, Schafer and Feiss, Mimura). Histological proof of the regeneration of the synapse in vago-sympathetic anastomoses has been furnished by

* Since the papers by Langley are very well known they have not been included in the references.

De Castro (1930, 1934, 1936–7) and Baron (1934). The regenerated synaptic endings in these cases are very similar to those present in the intact superior cervical ganglion. This is not surprising since the vagus carries numerous preganglionic fibres. In anastomoses of the cervical sympathetic trunk with the hypoglossus (De Castro, 1934, 1936–7), or with the phrenic nerve (Baron), the fibres of these nerves —normally supplying the musculature of the tongue and the diaphragm respectively—grow into the cervical trunk and reach the superior cervical ganglion, where they come into synaptic contact with its neurones. Since motor fibres are usually thick the synaptic endings are coarser than those of the preganglionic fibres but, as shown by De Castro, they end as rings and small reticulated enlargements. That these synapses are functional has been demonstrated by the two authors mentioned.

The effects of section of nerves which carry preganglionic fibres and their subsequent regeneration as evidenced in functional recovery, on the one hand, and the data gathered from histological studies of the same ganglia and nerves at various times after operation, on the other, are in such close agreement that their importance cannot be dismissed lightly. Gradual loss of function of the ganglion goes *pari passu* with the progressive breakdown of the synaptic terminals and preganglionic fibres. Conversely, the reappearance of functional activity coincides with the regeneration of new preganglionic nerve endings effecting synaptic contact with the dendrites and cell bodies or perikarya of the neurones of the ganglion.

(4) *Absence of transneuronal effect on synaptic endings during chromatolysis of the underlying neurone.* As is well known, most neurones swell and their Nissl bodies tend to disappear (chromatolysis) after their axons have been severed. The effect is transitory and is followed by recovery of the neurones. Barnard (1940) reported that section of the ventral roots of spinal nerves results in early disappearance of the *boutons* which end on the motor neurones of the ventral horn of the spinal cord. This has not been confirmed by Barr (1940) and Schadewald (1940). The latter author in a more recent article (1942) has shown that resection of several nerves (sciatic, femoral, trochlear and abducens) in newborn kittens neither alters the time of appearance nor the number of *boutons* developing about somatic motor neurones whose axons have been severed. Similarly Acheson, Lee & Morison (1942) have show that section of the phrenic nerve causes no detectable disturbance in *bouton* arrangement in the phrenic nucleus within 220 days. They have also demonstrated that loss of function occurs only during the chromatolytic changes and recovery of the neurones on which the *boutons* end. The absence of significant changes in the synaptic endings in the presence of chromatolysis does not, therefore, support the view of neurofibrillar continuity at the synapse.

(5) *Physiological evidence.* The coincidence of resumption of functional activity in a ganglion and the regeneration of the synaptic endings in contact with its neurones adds significance to the synapses described by histologists. Physiologists have long maintained that synaptic endings are the site of discharge of the nervous impulse, and in recent times evidence has been presented which indicates that synaptic transmission is effected by liberation of acetylcholine at the endings. In any event there is a definite interval between discharge of the impulse (or acetylcholine) at the *boutons* and the excitation of the underlying neurone; this interval (synaptic delay) in ganglionic synapses amounts to 2–4 milliseconds (Eccles, 1937); in the motor neurones of the brain stem it is still shorter (from 0·5–0·6 to 0·8–0·9 msec., Lorente de Nó, 1939). The synaptic delay thus varies within very small limits. Its existence cannot be accounted for if it be assumed that the neurofibrils of the preganglionic fibres pass without interruption into the dendrites and cell bodies of the ganglion cells, as has been claimed by the reticularists.

The action of certain drugs in establishing blockage of impulses reaching an autonomic ganglion has also been emphasized (Cajal, De Castro, Lawrentjew, Fedorow) as proof of the independence of the neurones in the autonomic nervous system. Langley found that nicotine paralysis does not prevent responses to postganglionic nerve stimulation. In a similar way a fibre passing through a particular ganglion without effecting synaptic contact with its neurones is not affected by application of the drug to that ganglion. If, on the contrary, the fibre is in synaptic association with the ganglion cells the transmission of impulses from the endings to the underlying neurone is prevented. This well-known property of nicotine—repeatedly verified by students in physiological laboratories—has been questioned by Stöhr as an argument in favour of the neurone theory. In one of his papers (1937) this author states that one cannot carry on work in histology by investigations with nicotine ('mit Nikotinversuchen kann man keine Histologie treiben', p. 365), an attitude of mind that does not require further comment.

Further physiological proofs of the independence of the neurones in the central as well as in the peripheral nervous system could be easily adduced because they abound in the literature. Indeed, had not the neurone theory been accepted by the physiologists the advances in the field of neurophysiology during the last fifty years would not have been possible. Bethe's view (1938) that many physiological results can be explained on the basis of the existence of a neurosyncytium is not likely to be shared by many neurophysiologists.

IV. THE SUPPOSED SYNCYTIAL ARRANGEMENT OF THE PERIPHERAL EXPANSIONS OF THE AUTONOMIC NERVOUS SYSTEM

The most persistent criticism of the neurone theory is to be found in a series of contributions on the structure of the peripheral portions of the autonomic nervous system appearing in the course of the last twelve years. As Levi (1941) has pointed out the publications on the subject agree on only one point, namely on the affirmation expressed in an apodictic manner that henceforth it will be necessary to renounce the neurone concept. Aside from this common point of view the findings of the reticularists disagree to such an extent that they will have to be considered separately.

(1) *Boeke's sympathetic 'ground plexus'*. We shall take up first Boeke's ideas on the structure of the sympathetic nerves and their mode of termination in diverse parts of the body, because his findings do not depart as widely from the older concepts as those of Stöhr and his students. The task is facilitated by the recent publication by Boeke (1940) of a book entitled *Problems of Nervous Anatomy* in which he presents in a condensed form his views on the subject. The importance he attaches to the structure of the autonomic nervous system is revealed in his statement that 'the whole battle concerning the structure of the nervous system and the connexions of its elements with each other and with the surrounding tissues has to be fought again over this peripheral system of nerve fibres' (p. 46). According to Boeke the finer sympathetic branches form a plexus of very fine, delicate, non-myelinated nerve fibres running in small bundles containing intercalated nuclei. In the coats of the human eyeball 'these bundles consist of very delicate varicose nerve-fibres and have such a peculiar appearance that they can be distinguished with the greatest certainty in every case from connective tissue surrounding them. They spring from the thicker nerve-bundles, to which they could always be traced. They do, however, follow an independent course through the connective tissue. The varicose nerve-fibres continuously anastomose with each other....The strands are often so loosely built, and the thin neurofibrillae run so far apart from each other although within the same protoplasmic formation, that an artificial adherence between them seems to be out of the question. The anastomoses visible must have been present in the living tissue. These loose strands of neurofibrillae are seen following the blood vessels everywhere. They run so close to the endothelium as sometimes to appear to be embedded in it; indeed, the neurofibrillae often lie in the same plane as the nuclei of the endothelial cells. Even capillaries are encircled by them. Besides the vascular bundles, we see similar strands running everywhere in the connective tissue. In the coats of the human eye, it is very easy to trace them to their origin from thicker nerve bundles, so that their nervous nature cannot be doubted' (p. 65). The general occurrence of a similar plexus in other locations, as, for instance, in the walls of small arteries and capillaries, in glands as the parotid and lachrymal, in the human skin and orbit and tongue of the hedgehog, cat and rabbit, is emphasized. 'It is found everywhere, and in favourable places in the impregnated preparations its connexion with the sympathetic ganglia is easily seen, as for instance in the musculature of the tongue, where sympathetic ganglia are scattered throughout the interstices of the muscle-bundles and therefore often lie in the neighbourhood of the blood-vessels studied' (p. 68).

The views of Boeke have been the subject of much criticism, and the structures so profusely represented in his papers have been interpreted in various ways. Hinsey (1934) 'is convinced that it isn't that other people have not seen the so-called "ground plexus" but rather that they have been unwilling to call structures of this kind nervous. The Bielschowsky stain is well-known to stain connective tissue and in deeply stained preparations this appears black. Especially where the fibres are small, as in reticular connective tissue, networks of this kind may be seen' (p. 554). I share this belief and have shown (Nonidez (1937*b*)) that the networks of the 'ground plexus' are not neurofibrillar but actually composed of argyrophil connective tissue fibres occurring among the nerve fibres of the bundle, as demonstrated with silver carbonate, which does not impregnate nerve fibres. The argyrophil connective tissue fibres are stained with the ammoniacal silver oxide solution of the Bielschowsky technique and cannot be differentiated from the impregnated axons. Nageotte (1938) does not believe that the reticulated bands are composed of neurofibrils but that they are actually bundles of axons within strands of the Schwannian syncytium; he explains the aspect of the bands on the assumption that water and other fluids are extracted from the nerve fibres during fixation, and that these fluids gather into drops which displace the axons and cause them to appear wavy and to touch each other as if they were anastomosed. Pensa (1937) claims that the method of Bielschowsky tends to separate the neurofibrils of the axons in many places along the fibres, and considers the reticulated aspect of the ground plexus (as well as the 'terminal' reticulum of Stöhr) as an artifact. Levi (1935) and Stefanelli (1938) agree in the main with Pensa. Finally, Schimert (1938*a*) in sympathetic ganglia transplanted into the peripheral segment of a previously sectioned ischiadic nerve has observed a structure identical with the ground plexus, but he regards it as a primitive growth form of the peripheral nerve fibres instead of being a terminal formation as claimed by Boeke.

In a paper answering my criticism Boeke (1938*b*) claims that the nervous nature of the reticulated bands has been demonstrated in the frog's tongue with the methylene-blue technique by his student Leeuwe (1937), and he copies one of the figures of

this author. It is obvious, however, that what Leeuwe has seen is not a neurofibrillar meshwork but a plexus of axons twisted and so displaced as to touch each other at many points, because methylene blue *does not* stain neurofibrils. Abrahám (1938), with the same material and technique, maintains that the axons of the plexus do not anastomose. Boeke further insists that a trained observer can easily differentiate neurofibrils from argyrophil connective tissue fibres. Such a distinction is often impossible in the case of the finest argyrophil meshworks. In fact, in preparations impregnated with silver carbonate I have shown (Nonidez 1936*b*, 1937*a*, *b*) that the structures revealed are identical with those regarded by Boeke as nerve bands. The occurrence of well-impregnated argyrophil networks in the walls of the blood vessels, and in their immediate vicinity in block impregnations with the Bielschowsky technique, I have explained as the result of incomplete impregnation of the continuous argyrophil connective tissue reticulum. 'Silver carbonate, the tannin silver technique and the Bielschowsky ammoniacal silver oxide solution yield meagre results when applied to blocks, the impregnation being irregular and often restricted to the periphery and the vicinity of the vessels within the block' (Nonidez, 1937*a*, p. 11).

A convincing demonstration of the nervous nature of the ground plexus will require (1) its impregnation with a technique that will not impregnate the argyrophil connective tissue fibres; this excludes the method of Bielschowsky and other procedures requiring the use of ammoniacal silver oxide solutions (with such neurofibrillar methods as those of Cajal, Bodian, Davenport, etc., reticulated bands have not been demonstrated in any of the organs investigated); (2) degeneration of the plexus after destruction of the sympathetic ganglia, sympathetic chain or section of sympathetic postganglionic nerves. Until this has been done we have a right to adhere to the older views which claim that the axons of nerve cells and their branches do not normally anastomose or break up into networks, but that they may form more or less elaborate plexuses in which twisting and frequent crossing of the axons simulates a network.

Continuity of the neurofibrils of the nerve twigs with the cytoplasm of the innervated cells has also been long maintained by Boeke. The existence of a 'periterminal network' in the motor plates of skeletal muscle has been denied by several competent histologists. Hinsey (1934) gives a critical review. Similar structures in cardiac and smooth muscle fibres, gland cells, etc., are also highly questionable. In the case of the heart Boeke (1933) has described a 'ground plexus' among the cardiac muscle fibres. Since he had previously shown that the nerve fibres ending in the myocardium terminate as hypolemmal loops he concludes that there are two types of innervation of the myocardium, and he further suggests that the fibres ending in loops may be the vagus postganglionics, whereas the sympathetic fibres pass into the 'ground plexus'. As to the existence of the 'periterminal networks' described by Boeke in the cardiac muscle fibres, Schimert (1937) claims that these structures are not continuous with the neurofibrils since they persist intact after degeneration of the nerve fibres. I have found no true hypolemmal nerve endings of the cardiac parasympathetic postganglionics which, with a modification of the Cajal silver technique, are sharply impregnated and can be differentiated from the faintly stained sympathetic postganglionics (Nonidez, 1939).

(2) *The 'terminal' reticulum.* Stöhr and his students (Reiser, Sunder-Plassmann, Seto) not only believe in the existence of anastomoses of the finest nerve branches but affirm that the ganglion cells from which they arise are connected with each other by means of their processes. The extra-ganglionic or peripheral processes merge into a delicate meshwork which forms an intermediate protoplasmic substance surrounding every cell of the body (including connective tissue and fat cells) and enters its cytoplasm. On account of this the finest meshworks have been designated collectively as the 'terminal' reticulum, whilst the reticulated bands leading to them have been called 'preterminal' reticulum (Reiser). There are no nerve endings: the structures described as endings by other authors are simply regarded as artifacts due to incomplete impregnation of the neurofibrils. In view of this fact, and of the continuity of the neurofibrils with the supplied cells, the term 'terminal' reticulum is most inappropriate. Scattered in the reticulum there are nuclei of the cells of Schwann, but in the more recent papers their existence has been denied.

The presence of a 'terminal' reticulum has been reported in diverse organs prepared with the Bielschowsky-Gros technique applied to frozen sections. It was first seen in the walls of the blood vessels by Reiser (1932). Since that time a rather voluminous literature has accumulated in which the structure under consideration has been described in the alimentary tract by Stöhr (1932, 1934, 1937), Reiser (1932, 1935*a*) and Ottaviani (1937); in the blood vessels by Stöhr (1935*a*, *b*, 1937, 1938); in the cardiac paraganglia and heart by Seto (1935, 1936); in the adrenal gland by Stöhr (1935*b*) and Sunder-Plassmann (1935); in the thyroid by Sunder-Plassmann (1934, 1935); in the pressoreceptor area of the carotid and in the glomus caroticum by the same author (1933); in the eye by Reiser (1935*b*, 1936) and Borri (1939); in the skin by John (1939); and in the respiratory system by Hayashi (1937) and Sunder-Plassmann (1938). Reviews have been issued by Reiser (1933), Stöhr (1935*a*, 1939*c*) and Rossi (1939). In its most extreme form, the theory of the 'terminal' reticulum claims the presence of this network in continuity with afferent or sensory nerve endings (Sunder-Plassmann, 1933; Stöhr, 1939*b*). Its proponents also speak of a merging of the terminations of the parasympathetic and sympathetic system, respectively,

into a common meshwork, and some, as for instance Reiser, go as far as supposing that all nerves, efferent as well as afferent, pass into this meshwork. The nerves of the body would, therefore, ultimately reach a vast nervous syncytium which is located in every interstice of the organs and even supplies the connective tissues and fat cells.

In the opinion of Stöhr (1938), Seto (1936) has demonstrated in the heart with particular clearness the merging of the vagus and sympathetic nerve fibres into a delicate 'terminal' reticulum. Years before, Lawrentjew (1929 *a*) working with the same material and technique failed to see the reticulum, or at least does not mention it as a nervous structure. Schimert (1937), also with the Bielschowsky-Gros method, has often seen reticulated structures in his preparations of the heart, but he could never confirm their continuity with the nerve fibres; more often the networks were seen in relation with argyrophil connective tissue fibres. More recently, using one of the formulae of the Cajal technique (fixation with alcohol-chloral hydrate), I have pointed out (Nonidez, 1939) differences in the staining of the parasympathetic (vagus) postganglionics and the sympathetic postganglionics; the latter stain faintly while the parasympathetic fibres are strongly impregnated, a difference which has been confirmed experimentally (Nonidez & Hare, 1942). The parasympathetic postganglionics have definite nerve endings appearing as minute clubs, rings and reticulated enlargements. The views of Stöhr and his students have received confirmation in the work of Stefanelli (1938) with the gold chloride technique of Ruffini, but he believes that a continuous 'terminal' reticulum (or 'diffuse' reticulum, as he calls it) is present only in the cutaneous expansions of the parasympathetic system whereas in the sympathetic the reticulum is discontinuous. The value of the gold chloride technique for the demonstration of the finest nervous structures is open to serious question. Rossi (1939), who has extensively reviewed the findings of Boeke and Stöhr, also favours the views of the last-named author.

Outside of Germany, Italy and Japan the theory of the 'terminal' reticulum has not found adherents. Specific denials of its presence in organs in which it has been described are found in the papers of Nonidez (1936 *b*, 1937 *a*) on the innervation of the blood vessels and of the thyroid, Schimert (1937) and Nonidez (1939) in the heart, Nonidez (1936 *a*) and Hollinshead (1939) in the supracardial and the aortic paraganglia, and Hollinshead (1936), Swinyard (1937), and McFarland & Davenport (1941) in the adrenal gland. Langworthy & Murphy (1939) and Alexander (1940) also deny the presence of the structure under consideration in the urinary bladder and the biliary system, respectively. As early as 1933 Boeke suggested that the 'terminal' reticulum looks more like a meshwork of connective tissue fibres than like a nervous structure, a belief shared by myself, having been able to demonstrate the existence of identical delicate meshworks with silver carbonate, which does not impregnate nerve fibres and their branches (Nonidez, 1936 *b*, 1937 *a*, *b*). Lawrentjew and his students, who have used the Bielschowsky-Gros technique extensively, have impregnated the 'terminal' reticulum in sections of organs fixed with strong formalin solutions (30–40 %) and, accordingly, believe that its impregnation with silver is due to overfixation, which does not take place when weaker formalin solutions are employed.

As pointed out by Gibson (1940) the claims of the reticularists are not backed by experimental evidence. A few degeneration experiments showing conclusively that the 'terminal' reticulum disappears after section of the nerves with which it is connected would be far more convincing than volumes of descriptions of this structure in intact organs. This is an important point that the reticularists in general seem to have ignored in their efforts to supplant the neurone theory with a concept which they believe more in accordance with the facts. The only attempt toward experimental verification of the nervous nature of the 'terminal' reticulum has been made by Reiser (1937), who, after extirpation of the Gasserian ganglion found degeneration of the fibres in the larger nerve bundles of the cornea; the 'terminal' reticulum, however, *remained intact*. This result could be anticipated if it is not a nervous structure.

V. CONCLUSION

Our present conception of the structure of the nervous system is not merely the result of histological observations but has been gradually built up with evidence derived from the closely allied fields of experimental anatomy, embryology, physiology, pathology and pharmacology. It would be surprising, therefore, to find that a single histological procedure could invalidate the data accumulated during many years of painstaking research with different methods of approach. Furthermore, the Bielschowsky technique, which according to Boeke, and to Stöhr and his school, reveals the syncytial structure of the peripheral nervous system in a clean cut fashion, does not yield the same results in the hands of other equally expert investigators. The researches of Lawrentjew and his students, Schimert and others, show that the same procedure used by Stöhr gives images which are practically identical with those obtained with other neurofibrillar methods.

Aside from the fact that, as already stated, the method of Bielschowsky often impregnates the argyrophil connective tissue fibres and, therefore, is not a technique as specific for nervous structures as other silver impregnation methods, the position of the reticularists is further weakened by lack of agreement on the interpretation of their findings. For several years Stöhr and his students have claimed that the 'ground plexus' and the 'terminal' reticulum are identical structures. This claim, how-

ever, has been repeatedly rejected by Boeke, who believes that the 'terminal' reticulum is largely composed of connective tissue fibres. Finally, Reiser (1935*b*) has been compelled to admit that the 'ground plexus' and the 'terminal' reticulum are morphologically two fundamentally different nervous structures (*morphologisch grundverschiedene Nervenformationen*). Since the reticularists cannot agree among themselves in offering a substitute for the neurone theory it would seem wise to retain it—as a working hypothesis at least—until better proof is offered of the syncytial nature of the nervous system.

VI. SUMMARY

During the last decade certain authors have considered the classical neurone theory to be obsolete and have resuscitated the long-rejected hypothesis of a nervous syncytium, mainly on histological evidence. Bauer believes in connexions between embryonic neuroblasts, serving as pathways for the growth of nerve fibres which thus develop intraprotoplasmically. This concept must be regarded as entirely hypothetical. Anastomoses found between neuroblasts and nerve fibres growing in tissue culture are emphasized by the reticularists, but this is not overwhelmingly against the neurone theory since such anastomoses are by no means universal and they depend on the part of the nervous system cultured and on the properties of the medium. Claims by Bauer on the continuity of nerve cells with each other and with neuroglia disagree with the demonstrated existence of free dendrites and of synapses with neurones. Reticularists also maintain the continuity of neurones in the autonomic system, but this too is open to serious criticism on histological grounds. Moreover, autonomic ganglia afford good material for the demonstration of synaptic endings both in fixed and living preparations. The reality of independent neurones is confirmed by the degeneration of synaptic *boutons* in the central nervous system after the destruction of nerve tracts, and by the degeneration and regeneration of preganglionic fibres in autonomic ganglia, in which the loss and recovery of function are confirmed experimentally. The existence of the synapse in confirmed by the delay in impulse transmission, varying with the type of synapse, and by the nicotine block to impulses from preganglionic fibres to autonomic ganglia. The recent criticisms of the neurone theory are based on preparations made exclusively by the Bielschowsky technique, unsupported by experimental work, and the several workers disagree widely in their evidence. Boeke claims that the finer divisions of the autonomic system contain a meshwork of neurofibrils clearly distinguishable from the surrounding connective tissue; this structure is not found, however, with other neurofibrillar methods but appears with silver carbonate which does not stain nervous structures. Stöhr believes in a fine terminal reticulum of the autonomic and sensory systems surrounding every cell in the body. Again other competent histologists, working with techniques which do not stain connective tissue fibres, have failed to confirm Stöhr, whose claims are moreover unsupported experimentally. The reticularists cannot agree among themselves on their substitute for the neurone theory and it would seem wise to retain this concept until better proof is offered of the syncytial nature of the nervous system.

VII. REFERENCES

ABRAHÁM, A. (1938): *Z. Zellforsch.* **27**, 745.

ACHESON, G. H., LEE, E. S. & MORISON, R. S. (1942): *J. Neurophysiol.* **5**, 269.

ALEXANDER, W. F. (1940): *J. Comp. Neurol.* **72**, 357.

ANGULO, A. W. (1939): *J. Comp. Neurol.* **71**, 325.

BALADO, M. & FRANKE, E. (1930): *Arch. Argent. Neurol.* **6**, 77.

BARNARD, R. I. (1940): *J. Comp. Neurol.* **73**, 235.

BARON, M. (1934): *Z. mikr.-anat. Forsch.* **35**, 331.

BARR, M. L. (1939–40): *J. Anat.* **74**, 1. — (1940): *Anat. Rec.* **77**, 367.

BARRON, D. H. (1943): *J. Comp. Neurol.* **78**, 1.

BARTELMEZ, G. W. and HOERR, N. L. (1933): *J. Comp. Neurol.* **57**, 401.

BAUER, K. (1932): *Z. mikr.-anat. Forsch.* **28**, 47. — (1936): **39**, 57. (1938): **43**, 48. — (1940): *Z. Zellforsch.* **30**, 751.

BETHE, A. (1938): *Verh. Anat. Ges.*, 45 *Vers., Koenigsberg, Anat. Anz.* **85**, 142.

BODIAN, D. (1937): *J. Comp. Neurol.* **68**, 117. — (1940): **73**, 323. — (1942): *Physiol. Rev.* **22**, 146.

BOEKE, J. (1933*a*): *Z. mikr.-anat. Forsch.* **33**, 233, 276. — (1933*b*): **34**, 330. — (1934*a*): **35**, 551. — (1934*b*): *Bull. Histol. Tech. micr.* **11**, 49, 221; **13**, 113. — (1935): *Zeit. mikr.-anat. Forsch.* **38**, 554, 594. — (1936): **39**, 477. — (1938*a*): *Verh. anat. Ges.*, 45 *Vers. Koenigsberg, Anat. Anz.* **85**, 111. — (1938*b*): *Anat. Anz.* **86**, 150. — (1940): *Problems of Nervous Anatomy.* Oxford.

BORRI, N. (1939); *Z. Zellforsch.* **29**, 128.

CAJAL, S. R. (1934): *Trab. Lab. Invest. biol. Madr.* **29**, 1. — (1935): Bumke and Foerster's *Handbuch der Neurologie*, **1**. Berlin.

CLARK, W. E. LE GROS & PENMAN, G. G. (1934): *Proc. Roy. Soc.* B, **114**, 291.

DE CASTRO, F. (1922): *Trab. Lab. Invest. biol. Madr.* **20**, 113. — (1930): **26**, 357. — (1932): Penfield's *Cytology and Cellular Pathology of the Nervous System*, **1**, 319. New York. — (1934): *Trab. Lab. Invest. biol. Madr.* **29**, 397. — (1936–7): **31**, 271.

DOLGO-SABUROFF, B. (1936): *Z. Anat.* **105**, 79.

ECCLES, J. C. (1937): *Physiol. Rev.* **17**, 538.

ESAKI, S. (1929): *C.R. Ass. Anat.* **24**. Bordeaux.

FEDOROW, B. G. (1935): *Trab. Lab. Invest. biol. Madr.* **30**, 403.

FEDOROW, B. G. & MATWEJEWA, S. J. (1935): *Trab. Lab. Invest. biol. Madr.* **30**, 379. — (1936–7): **31**, 203.

FOERSTER, O., GAGEL, O. & SHEEHAN, D. (1933): *Z. Anat.* **101**, 553.

GIBSON, W. C. (1937): *Arch. Neurol. Psychiat.* **38**, 1145. — (1940): *J. Neurophysiol.* **3**, 237.

GLEES, P. (1941): *J. Anat., Lond.*, **75**, 434.

GLEES, P. & CLARK, W. E. LE GROS (1941): *J. Anat., Lond.*, **75**, 295.

GRIGORJEFF, L. M. (1931): *Arch. exp. Zellforsch.* **11**, 483. — (1933): **13**, 195.

HARTING, K. (1930): *Z. Zellforsch.* **12**, 518. — (1934): *Z. mikr.-anat. Forsch.* **35**, 631.

HAYASHI, S. (1937): *J. Orient. Med.* **27**, 37.

HELD, H. (1929): *Fortschr. naturwiss. Forsch.* Neue Folge, **8**.

HERZOG, F. & GÜNTHER, B. (1938): *Z. Neurol.* **160**, 550.

HIŇSEY, J. C. (1934): *Physiol. Rev.* **14**, 514.

HINSEY, J. C., PHILLIPS, R. A. & HARE, K. (1939): *Amer. J. Physiol.* **126**, 534.

HOFF, E. C. (1932*a*): *Proc. Roy. Soc.* B, **111**, 175. — (1932*b*): **111**, 226. — (1935): *Arch. Neurol. Psychiat.* **33**, 687.
HOFF, E. C. & HOFF, H. E. (1934): *Brain*, **57**, 454.
HOLLINSHEAD, W. H. (1936): *J. Comp. Neurol.* **64**, 449. — (1939): **71**, 417.
HUZELLA, T. (1938): *Verh. Anat. Ges.*, 45 *Vers. Koenigsberg, Anat. Anz.* **85**, 91.
IWANOW, I. F. & RADOSTINA, T. N. (1932–3): *Trab. Lab. Invest. biol. Madr.* **28**, 303.
JOHN, F. (1939): *Z. Zellforsch.* **30**, 297.
JUBA, A. (1937): *Z. Zellforsch.* **26**, 396.
KIMMEL, D. L. (1940): *J. Comp. Neurol.* **72**, 83.
KOLOSSOW, N. G. (1932–3): *Trab. Lab. Invest. biol. Madr.* **28**, 345.
KOLOSSOW, N. G. & POLYKARPOWA, G. A. (1935): *Z. Anat.* **104**, 716. — (1936): **106**, 98.
KOLOSSOW, N. G. & SABUSSOW, G. H. (1932): *Z. mikr.-anat. Forsch.* **29**, 541.
KOLOSSOW, N. G., SABUSSOW, G. H. & IWANOW, J. F. (1932): *Z. mikr.-anat. Forsch.* **30**, 257.
LANGWORTHY, O. R. & MURPHY, E. L. (1939): *J. Comp. Neurol.* **71**, 487.
LAWRENTJEW, B. J. (1924): *Anat. Anz.* **58**, 529. — (1925): *Z. mikr.-anat. Forsch.* **2**, 201. — (1929*a*): **16**, 383. — (1929*b*): **18**, 233. — (1930): **23**, 527. — (1934): **35**, 71.
LAWRENTJEW, B. J. & BOROWSKAJA, A. J. (1935): *Z. Zellforsch.* **23**, 761.
LAWRENTJEW, B. J. & NAIDITSCH, M. S. (1932–3): *Trab. Lab. Invest. biol. Madr.* **28**, 223.
LAZARENKO, TH. (1931): *Arch. exp. Zellforsch.* **11**, 553.
LEEUWE, A. (1937): Inaug. Dissert. Utrecht.
LEVI, G. (1934): *Z. ges. Anat.* 3. *Ergebn. Anat. EntwGesch.* **31**, 125. — (1935): *Trattato d'Istologia, Torino.* — (1941): *Arch. Biol.* **52**, 133.
LEVI, G. & DELORENZI, E. (1934): *Arch. Ital. Anat.* **33**, 443.
LEWIS, W. H. & LEWIS, M. R. (1912): *Anat. Rec.* **6**, 7.
LORENTE DE NÓ, R. (1938): *J. Neurophysiol.* **1**, 195. — (1939): **2**, 402.
MCFARLAND, W. E. & DAVENPORT, H. A. (1941): *J. Comp. Neurol.* **75**, 219.
MAKSUDOWA, M. A. (1936): *Anat. Anz.* **83**, 19.
MEYER, H. & JABLONSKI, W. (1937): *J. Anat., Lond.*, **72**, 62.
MIHÁLIK, P. (1932): *Anat. Rec.* **54**, 149. — (1935): *Arch. exp. Zellforsch.* **17**, 119. — (1940): *Anat. Anz.* **89**, 241.
MINCKLER, J. (1940): *Anat. Rec.* **76**, 42; **77**, 9. — (1941): *Arch. Neurol. Psychiat.* **45**, 44. — (1942): *Amer. J. Pathol.* **18**, 1061.
MINKOWSKI, M. (1920): *Schweiz. Arch. Neurol. Psychiat.* **6**, 201.
NAGEOTTE, J. (1938): *Anat. Anz.* **87**, 49.
NONIDEZ, J. F. (1936*a*): *J. Anat.* **70**, 215. — (1936*b*): *Anat. Anz.* **82**, 348. — (1937*a*): **84**, 1. — (1937*b*): **84**, 315. — (1939): *Amer. J. Anat.* **65**, 361.
NONIDEZ, J. F. & HARE, K. (1942): *J. Comp. Neurol.* **76**, 91.
OTTAVIANI, G. (1937): *Z. Zellforsch.* **27**, 393.
PENSA, A. (1937): *Monit. Zool. Ital.* **47**, 217.
PHALEN, G. S. & DAVENPORT, H. A. (1937–8): *J. Comp. Neurol.* **68**, 67.
POLYÁK, S. L. (1941): *The Retina.* Chicago.
REISER, K. A. (1932): *Z. Zellforsch.* **15**, 761. — (1933): **17**, 610. — (1935*a*): **22**, 675. — (1935*b*): *Arch. Augenheilk.* **109**, 251. — (1936): **109**, 481. — (1937): **110**, 253.
ROMANES, G. J. (1941): *J. Anat., Lond.*, **76**, 112.
ROSSI, F. (1939): *Riv. Patol. nerv.* **53**, 216.
SAKAGUCHI, Z. (1939): *J. Orient. Med.* **30**, 247, 251.
SCHADEWALD, M. (1940): *Anat. Rec.* **76**, 48. — (1942): *J. Comp. Neurol.* **77**, 739.
SCHIMERT, J. (1937): *Z. Zellforsch.* **27**, 246. — (1937–8): *Z. Anat.* **108**, 761. — (1938*a*): *Verh. Anat. Ges.*, 45 *Vers. Koenigsberg, Anat. Anz.* **85**, 173. — (1938*b*): *Z. mikr.-anat. Forsch.* **44**, 85. — (1939*a*): *Verh. anat. Ges.*, 46 *Vers. Leipzig, Anat. Anz.* **87**, 312. — (1939*b*): *Z. Anat.* **109**, 665.
SCHROEDER, P. (1938): *Verh. anat. Ges.*, 45 *Vers. Koenigsberg, Anat. Anz.* **85**, 153.
SEREBRJAKOW, P. (1929): *Z. Zellforsch.* **9**, 425. — (1934): **22**, 140. — (1936): **24**, 152.
SETO, H. (1935): *Z. Zellforsch.* **22**, 213. — (1936): *Arb. anat. Inst. Sendai*, **19**, 1.
SILVER, M. (1942): *J. Comp. Neurol.* **77**, 1.
SNIDER, R. S. (1936): *J. Comp. Neurol.* **64**, 417.
STEFANELLI, A. (1938): *Z. Zellforsch.* **28**, 485.
STÖHR, P. (1932): *Z. Zellforsch.* **16**, 123. — (1934): **21**, 243. — (1935*a*): *Z. Anat.* **104**, 133. — (1935*b*): **104**, 475. — (1937): *Z. Zellforsch.* **27**, 341. — (1938): *Z. ges. Anat.* 3. *Ergebn. Anat. EntwGesch.* **32**, 1. — (1939*a*): *Z. Zellforsch.* **29**, 569. — (1939*b*): **30**, 78. — (1939*c*): *Klin. Wschr.* **18**, 41.
SUNDER-PLASSMANN, P. (1930): *Z. Anat.* **93**, 567. — (1933): *Z. Neurol.* **147**, 414. — (1934): *Klin. Wschr.* **13**, 364. — (1935): *Dtsch. Z. Chir.* **245**, 756. — (1938): **250**, 705. — (1939): **252**, 1,210.
SWINYARD, C. A. (1937): *Anat. Rec.* **68**, 417.
SZANTROCH, Z. (1933*a*): *Arch. exp. Zellforsch.* **14**, 442. — (1933*b*): *Z. Zellforsch.* **20**, 417. — (1935): **23**, 464.
SZEPSENWOL, J. (1939): *Rev. Soc. argent. Biol.* **15**, 75.
TELLO, J. F. (1923): *Trab. Lab. Invest. biol. Madr.* **21**, 1.
TIEGS, O. W. (1927): *Austr. J. Exp. Biol. Med. Sci.* **4**, 25, 79. — (1931): *J. Comp. Neurol.* **52**, 189.
WEISS, P. (1934): *J. Exp. Zool.* **68**, 393.
WEISS, P. & WANG, H. (1936): *Anat. Rec.* **67**, 105.
WINDLE, W. F. (1932): *J. Comp. Neurol.* **55**, 315. — (1933): **58**, 643.
YAMASHITA, K. (1939): *J. Orient. Med.* **30**, 150, 153.

DIGESTION IN THE RUMINANT

By R. A. McANALLY and A. T. PHILLIPSON, Unit of Animal Physiology, Cambridge

(*Received 2 August 1943*)

I. INTRODUCTION

The alimentary canal of ruminants is specialized in ways which serve to digest the extremely refractory materials on which these herbivores live; the most remarkable development is the rumen—a diverticulum of the stomach which is in effect a fermentation chamber where cellulose and other insoluble articles of diet can be subjected to the action of free organisms. To understand the digestive processes of the ruminant it is necessary to consider the chemical nature of the food, the movements and physical environment to which it is subjected and the chemical processes responsible for its digestion.

Chemical composition of foodstuffs. The diet of the grazing ruminant in the wild state consists largely of grass at its various stages of growth, while for the domesticated ruminant grass is of importance more during the summer than during the winter months when, owing to slow rate of growth, grass must be supplemented by other foodstuffs such as hay, straw, cereals, roots, green vegetables and high protein meals and cakes. Of the three main 'proximate principles' of food, namely, carbohydrates, proteins and fats, the fats are of minor interest since their amount is small in the diet of the ruminant, except when large quantities of concentrates are given. There is little evidence to suggest that the rumen plays any part in their digestion. Protein is to some extent broken down in the rumen, but the organ is not essential for its digestion; however, bacterial synthesis of protein may occur from metabolic products of protein in the plant such as the amides asparagine and glutamine and the aminoacids; further, this fraction may be augmented by feeding with urea or ammonium salts. The chief sources of protein for the domestic ruminant are grass, particularly if it is close grazed or frequently cut, legumes and manufactured concentrates.

The most interesting group of foodstuffs from the point of view of rumen digestion is the carbohydrates, since all representatives of the group in plant materials are probably attacked in the rumen and the majority would be undigested were it not for this organ. Free sugars are found in roots, two-thirds of the dry matter of the mangold being sugar (Wood & Berry, 1905), but their amount is not significant in other fodders. The chief sources of starch are the cereals. Cellulose is by far the most important carbohydrate for herbivorous nutrition: it is present to a greater or lesser extent in all plant materials, and as will be shown its digestion takes place largely in the rumen. As much as 49 % of the dry matter of September-cut grass (Norman & Richardson, 1937) consists of 'crude cellulose' by which term Norman means chemically pure cellulose together with xylan—a polysaccharide which is made up of pentose units and which in the plant is associated with cellulose in an unknown manner. Wheat straw contains an even higher percentage of crude cellulose as defined above, namely 50 % (McAnally, 1942). The quantity in young grass is, however, much less: thus in a May-cut sample of the same species (rye grass) Norman found only 26 %. As the grass gets older the relatively indigestible substance lignin increases in amount; not only is lignin itself somewhat indigestible but it appears to prevent the digestion of a part of the cellulose. Therefore, although the older grass contains more cellulose than the younger, from the point of digestibility the disparity in cellulose available for nutrition is not so great as might be expected because of the refractory influence of the lignin.

The importance of other less well-known carbohydrates in ruminant nutrition should not be overlooked, because the amounts which may be ingested by an animal are far from negligible. Thus Norman & Richardson (1937) have shown that as much as 25·6 % of the dry matter of rye grass cut in May consisted of a fructosan. Phillips, Davis & Weihe (1942) found that 22·9 % of the dry matter of Timothy grass harvested in June was pentosans. Xylan associated with true cellulose in the plant constitutes 10–15 % of the dry matter of winter grass and straw (Norman, 1935, 1937; McAnally, 1942). Polyuronides in Norman's September rye grass were 8·6 % of the dry matter and Phillips *et al.* found this fraction to constitute 4·8 % of the dry matter of their Timothy grass.

The significance of alimentary micro-organisms in digestion. Since there are no known enzymes present in the alimentary secretions which are capable of hydrolysing substances such as cellulose, the presence of micro-organisms appears to be necessary for their digestion. Pasteur (1885) suggested that the presence of bacteria in the alimentary tract of the animal was essential for normal life of rabbits, guinea pigs, hens and dogs. This hypothesis was investigated by Nuttall & Thierfelder (1895–6) who

delivered foetuses from guinea-pigs at full term by Caesarian Section under strict aseptic conditions and fed them in sterile surroundings upon milk and biscuits. Under these conditions their animals lived up to thirteen days and put on weight. The absence of bacteria in the alimentary tract was proved subsequently by cultural methods. Cohendy & Wollman (1914) obtained similar results with guinea-pigs fed upon milk, sterilized hay, lucerne and bran but were able to maintain their animals up to twenty-nine days. Kuster (1914), however, produced evidence that cellulose was not digested by a bacteria-free kid goat, which lived for thirty-five days on a milk diet to which cellulose in the form of rye bread was added. Only negligible amounts, 1·15%, of the cellulose ingested was lost in its passage through the animal. Glimstedt (1936) repeated Nuttall's experiments with better results as the necessary vitamins and minerals were included in the diet but even so a heavy death rate occurred although some animals lived as long as sixty days in a bacteria-free condition.

These experiments demonstrate conclusively that the presence of bacteria in the alimentary tract is not essential for the early stages of life provided the necessary calories and proteins are present in a form which can be digested by the alimentary secretions. If, however, the animal relies upon cellulose and its associated carbohydrates for its source of calories, then it would appear that bacteria are essential for digestion as their absence would mean starvation. In this sense Pasteur's hypothesis may be accepted when applied to herbivorous animals feeding under natural conditions.

Similar experiments have been done with sheep and goats to determine whether the ciliates of the rumen play an essential role in digestion. Young suckling ruminants do not harbour these organisms which make their appearance in the rumen only when the animals start to eat substantial amounts of solid food. The ciliates do not, as far as is known, form cysts and the period of survival outside the rumen is short, although Fantham (1922) states that he was able to find *Entodinium* and *Diplodinium* on wet grass and in aqueous washings of fresh grass and even on dried grass from sheep runs in South Africa. In spite of this observation it is possible to maintain animals without ciliates after they have been freed of these organisms even though they are fed upon unsterilized food including green grass, provided they are isolated from infected animals and strict precautions are taken to prevent mechanical transmission from normal animals (Becker & Hsuing, 1929; Mangold & Radeff, 1930–1). It appears, therefore, that natural transmission occurs by contamination of food and water, probably with saliva, since neither free-living nor encysted ciliates have been found in the faeces. Feeding experiments with goats during (1) a period in which they were ciliate-free, and (2) after reinfection, have shown no significant difference in the digestibility coefficients of crude fibre, hemicellulose, pentosan, α-cellulose and protein (Becker, Schulz & Emmerson, 1929–30; Winogradow, Winogradowa-Fedorowa & Wereninow, 1930). Comparison of growth rates of ciliate-free and normally infected lambs carried out over a period of 19 weeks has also failed to reveal that ciliate-free animals are in any way handicapped, in fact the ciliate-free ruminants grew a little more rapidly than the controls (Becker & Everett, 1930; Polyansky & Strelkow, 1935). It is clear that the presence or absence of ciliates in the rumen has little effect upon the well-being of the animal, and it can be assumed that any role they may play in digestion is not essential.

The alimentary tract of herbivorous animals. Two methods of dealing with bulky fibrous foods have been evolved by herbivorous animals, either (1) the stomach (ruminants, e.g. cow, sheep, etc.) or (2) the large intestine (horse, rabbit) is enlarged and developed into a fermentation chamber.

Between the two extremes intermediate types exist. The hippopotamus and kangaroo have comparatively large stomachs which are sacculated along the greater curvature and which bear cardial diverticula. The large gut in most herbivorous and omnivorous animals is capacious compared with that of carnivores, although relatively it is less exaggerated than that of the horse. The stomach of those herbivores in which it is uncomplicated and of omnivores differs from that of the carnivores in that the cardial end is enlarged and in some cases (hamster) forms a distinct pouch. The histology of the stomach of these intermediate types is consistent with the possibility of fermentation taking place therein, for the cardial end is lined with stratified squamous epithelium (horse, rat, etc.) or else by glands devoid of central or oxyntic cells (pig).

The special development of the large gut and the stomach in the animals referred to above has the object of providing not only suitable conditions for fermentation but also a means of delaying the passage of food so that effective time for fermentation may occur. In the simple stomach delay is short, but the food is retained in the large gut, which on account of the sacculations of the walls of the caecum and colon efficiently delays the passage of food. The sacculations of the stomach of the hippopotamus and kangaroo presumably serve a similar purpose, but the stomach of the ruminant is still better equipped as the passage of exit is narrow and is guarded by a filter, the omasum. In the camel, in which the omasum is absent, sacculations again appear in the lower walls of the reticulum and colon and possibly compensate for the absence of the omasum.

II. THE MECHANICAL PROCESSES OF DIGESTION IN THE RUMINANT

The movements of the stomach have been studied in the cow (Wester, 1926; Schalk & Amadon, 1928) through large open rumen fistulae and in the sheep

and goat both by radiological observation and by pressure tracings (Czepa & Stigler, 1926, 1929; Krzywanek & Quast, 1937; Quin & van der Wath, 1938; Phillipson, 1939). The sequence of events is essentially the same. The *reticulum* and *rumen* are one functional unit as their movements are co-ordinated, and together they form one fermentation chamber of which the reticulum is the pacemaker.

Fig. 1 represents the cycle of movement found in the reticulum and rumen of sheep. The whole cycle represents 1 min. and is divided up into 5 sec. intervals. The degree of contraction is shown by the depth of shading; thus the height of contraction is represented by the regions shaded black. The outer circle illustrates the movements of the reticulum, the middle circle those of the dorsal sac of the rumen, and the inner circle those of the ventral sac of the rumen.

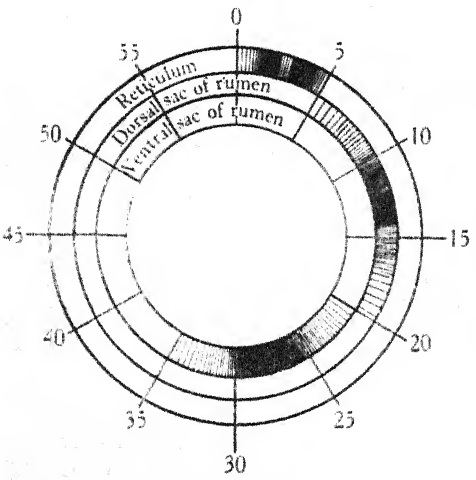


Fig. 1.

Different sheep yield tracings sometimes more complicated, sometimes less so, but they may be reduced to a common type of which the more complicated rhythms may be regarded as derivatives. The type is illustrated in Fig. 1; it involves three principal localities—the reticulum, the dorsal part of the rumen and the ventral part of the rumen. The whole cycle takes about 1 min. (*a*) We may commence with the reticulum because it is the pacemaker. The contraction of the reticulum, which takes about 5 sec., is a double contraction having therefore two peaks of which the later one is the higher. (*b*) Before the end of the fifth second the contraction of the dorsal part of the rumen commences and lasts approximately 15 sec., rising gradually to a maximal strength and fading at the same speed. (*c*) The contraction of the ventral sac follows that of the dorsal sac immediately and gives a similar curve lasting the same time. (*d*) The whole organ, reticulum and rumen, remains at rest from approximately the 35th sec. until the next contraction of the reticulum. Irregularities in this cycle are seen at two places: a pause may occur between the two contractions of the rumen (*b*) and (*c*), and the period of rest at the end of the cycle varies in length.

Tracings taken from the reticulum of the cow show two distinct contractions separated by complete relaxation, but in the sheep the second contraction is superimposed upon the first and relaxation between the two stages is exceptional. The movements of the rumen are slow and sustained compared with those of the reticulum, and the pressure variations are considerable, varying in the sheep from −6 to 94 mm. of water (Quin, van der Wath & Myburgh, 1938). The movements of the reticulo-rumen sac keep the ingesta in continual motion. The mass of ingesta of the reticulum presents radiologically a fluid level; when the organ contracts the liquid ingesta are thrown backwards from the reticulum to the anterior region of the rumen and spread over the solid ingesta in the dorsal sac. The anterior extremity of the dorsal sac of the rumen relaxes as the reticulum contracts, and Czepa & Stigler (1926) emphasize this alternate movement between the two organs resulting in an interchange of ingesta; they state that no correlation exists between these movements and those of the posterior sacs of the rumen —an observation which we have been unable to confirm. A study of tracings taken from the reticulum and posterior part of the rumen of both the cow and sheep indicate clearly that the movements of these organs are co-ordinate.

The *omasum* has been studied in respect of the movements visible radiologically, the pressure changes within it and the effect of food upon its development and structure. Conflicting observations regarding the movements (Czepa & Stigler, 1926, 1929; Magee, 1932; Phillipson, 1939) make it impossible to give a confident picture of its cycle at present. The position of the organ, however, changes during, and probably as the result of, the contraction of the reticulum. Changes in pressure within the lumen of the omasum of cattle have been recorded (Wester, 1926; Schalk & Amadon, 1928) and a fall occurs during the second contraction of the reticulum, the period in which the change in position is seen in lambs. The significance of this fall, however, is disputed. That the omasum has a mechanical function is indicated by the work of Trautmann (1933*b*) and Trautmann & Schmitt (1935), who point out that (1) the particles in the ingesta of the omasum near its entrance are coarser than those found near its exit, (2) if a young goat is fed entirely upon milk the omasum remains infantile and no development occurs until rough food is given. Similarly, if the laminae of the omasum are extirpated from older goats, regeneration, which occurs when normal food is given, is less noticeable if nothing but milk is given as food. The mechanical activity of the omasum, however, cannot be of great importance, for experi-

mental anastomosis of the cavities of the rumen and reticulum with the abomasum, so that the omasum is short-circuited, does not seriously impair the health or digestion of the animal so far as can be judged by appearance and rate of growth.

The *abomasum* is easy to examine radiologically, and all who have done so agree that little or no movement of the body of the stomach occurs in the normally filled organ; strong tonic contractions, however, can be seen in the abomasum of the hungry lamb (Phillipson, 1939). The body of this organ is subject to movement from external sources in the same way as the omasum; when the reticulum contracts, the fundus is lifted by means of the thin muscular connexions between these organs. This effect can be detected in older lambs (4 months or over) when the abomasum contains only small quantities of radio-opaque material. Movements of the wall of the organ are confined to the pyloric antrum where strong peristaltic waves occur.

Rumination. The earliest work on regurgitation which we need consider is that of Haubner (1837) and Flourens (1844). They both accepted the thesis that the regurgitated material consists of a bolus. This has been disproved by Colin (1886), who sutured the lips of the oesophageal groove so that only liquid could get through and found rumination still took place. A controversy, however, arose, as to how food gets into the oesophagus, then to be propelled by antiperistaltic movement up to the mouth. Granting that the regurgitated food is liquid, is it thrust upwards by contraction of the rumen or sucked upwards by a combination of inspiratory effort and closure of the glottis? Colin agreed with his predecessors that the ingesta are forced into the oesophagus by contraction of the rumen aided by contraction of the diaphragm, and he described the lower end of the oesophagus as forming an infundibulum to receive the ingesta. The theory of Colin dominated the literature until recently, overriding the alternative theory of Toussaint (1875), who, working under the direction of Chauveau, showed that during the act of regurgitation the animal made a deep inspiratory effort of the diaphragm with a closed glottis and produced a marked fall in pressure in the trachea by so doing which greatly exceeded the fall produced by normal inspiration. He was unable to find any contraction of the rumen at the moment of regurgitation, and recorded the passage of the ingesta through the lower cervical oesophagus immediately after the inspiratory effort. He attributed the passage of ingesta into the oesophagus to the negative pressure in the thorax. Colin rejected this theory, since neither section of the phrenic nerves nor insertion of a tracheal tube abolished rumination, but Toussaint, who was aware of these apparent weaknesses to his theory, considered that, as the animal had to use much greater effort to ruminate under these conditions, his theory was strengthened. The events described by Toussaint have been confirmed (Bergman & Dukes, 1926; Stigler, 1931), and the actual fall in pressure in the oesophagus itself during the inspiratory effort was recorded by Stigler who also showed that the slight stretching of the head that occurs during regurgitation effectively closes the cervical portion and seals the oesophagus from the outside air. Wester (1926) rejects Toussaint's theory as he considers the inspiratory effort insufficient to produce a flow of ingesta into the oesophagus and suggests that the oesophagus itself actively contracts in a longitudinal manner, and owing to its attachments to the stomach dilates and forms the infundibulum described by Colin. According to this theory the oesophagus is the organ mainly concerned in producing the fall in pressure, the inspiratory effort merely aiding it. The musculature of the oesophagus in ruminants is striated throughout and such an action is possible, but active contraction would make the walls more rigid and less dilatable. Consequently such an action would be antagonistic to the inspiratory effort, and, as it has been shown that the unimpaired fall of pressure in the thorax is essential for easy regurgitation, such an action of the oesophagus does not seem probable. Stigler (1931) found that unilateral pneumothorax prevented rumination in a goat, although repeated efforts were made to regurgitate; if the oesophagus was mainly responsible for the aspiratory action, then regurgitation should have been possible. The theory of Chauveau and Toussaint, therefore, appears to be correct.

The role of the intrinsic contraction of the oesophagus has been studied by radiological methods. In sheep and goats (Czepa & Stigler, 1929; Magee, 1932) the whole thoracic oesophagus becomes filled with radio-opaque fluid on regurgitation, and the upper part of the fluid is carried to the mouth by a wave of anti-peristalsis while the lower portion runs back to the rumen. An extra contraction of the reticulum occurs during rumination which raises the fluid level well above the cardia and precedes the normal movement, regurgitation occurring just after the peak of this contraction (Bergman & Dukes, 1926).

Passage of food through the ruminant alimentary tract. In the young ruminant milk suckled from the mother or from a bottle does not enter the reticulum and the rumen but passes directly through the omasum to the abomasum; this is achieved by the closure of the lips of the oesophageal groove. Closure of the groove does not occur when solid food is eaten; a bolus is formed which is deposited in the anterior region of the rumen (the vestibulum), and its subsequent destination depends to some extent on the specific gravity of the individual particles. The passage of food through the chambers of the stomach may be represented by the following scheme: From the oesophagus solid food may go: (1) To the reticulum, whence it can go (*a*) to the omasum, (*b*) to the rumen from which it is returned,

(*c*) back to the mouth through the oesophagus (regurgitation) when it may start afresh on either of routes 1 or 2. (2) To the rumen direct, from which it has only one escape, namely, to the reticulum.

The route taken depends to some extent on the nature of the food. The direct passage of milk via the omasum to the abomasum may be prolonged beyond the normal suckling age if the lamb is fed persistently from a bottle. Wise & Anderson (1939) found in calves that the act of suckling produced direct passage to the abomasum as both milk and water took this direct passage when taken from a bottle. However, milk drunk from a pail also in most cases passed directly to the abomasum, but water did not. Watson (1941) made a detailed analysis of all these four factors, milk, water, suckling and drinking, and his results suggest that the deciding

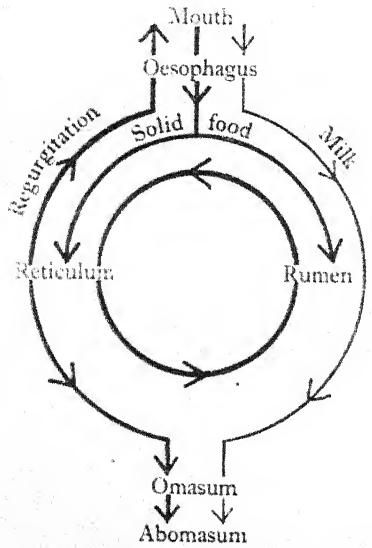


Fig. 2.

factor is psychological, for if a lamb wanted milk and thought it was being given milk direct passage to the abomasum took place regardless of which of the four factors were associated. If the lamb thought it was being given water to quench its thirst, then the fluid passed to the forestomach. The evidence points to the conclusion that closure of the groove is a conditioned reflex associated with milk. As between routes 1 and 2 from the oesophagus, heavy food such as maize (Schalk & Amadon, 1928) tends to sink into the reticulum whilst lighter foods, e.g. hay, go backwards to the rumen. Ritzman & Benedict (1938) have expanded the above observations into a general statement that all concentrated foodstuffs take the direct route from the reticulum to the omasum, thus avoiding the rumen. This generalization is not justified, as Nevens (1928) had previously shown that in a cow killed 6 hr. after eating 18 lb. of maize, 70 % of the corn was in the rumen, 3 % in the reticulum, 4 % in the omasum and 2 % in the abomasum; the remainder (21 %) was not accounted for.

The movements of food in the rumen are as follows. The ingesta which pass directly from the oesophagus to the rumen consist of the coarse food together with material washed back from the reticulum. In lambs taken from grass the ingesta in the rumen are seen radiologically to be heaped up, and a fluid level appears only in the later stages of digestion. The action of the dorsal sac is to compress the mass of ingesta so that the fluid is squeezed out to the ventral sac. Fine particles of food will, therefore, be washed through to the ventral sac more rapidly than large particles. The ingesta in the ventral sac are more fluid than those of the dorsal sac and the particles in a finer state of division. This is the result of the continual flushing of the ingesta in the dorsal sac with the liquid ingesta from the reticulum. As fermentation proceeds and the large particles in the dorsal sac are reduced in size, they pass to the ventral sac, whence they are moved forwards and upwards by the contractions of the ventral sac to the anterior region of the rumen from where they can be washed back to the reticulum. The three alternative routes, however, are still open to them, but should they be returned to the rumen either before or after remastication their subsequent passage will be more rapid owing to their reduction in size. The passage of any one meal through the rumen, therefore, is spread over a considerable period; this has been demonstrated by Usuelli (1933), who gave 1 kg. of stained oats to cattle, that were otherwise fed normally, and withdrew samples of ingesta from the rumen by stomach tube at regular intervals until no stained particles were left. In 2 days the concentration of stained oats in the rumen, estimated by weight of stained material per 100 c.c. rumen contents, had fallen by half, the decrease became progressively less and oats persisted in the rumen for 7 days. No comparable data exist for the small ruminants.

The actual mechanism which regulates the passage of ingesta from the reticulum to the *omasum* is not understood, in spite of the work of several observers (Schalk & Amadon, 1928; Wester, 1926; Phillipson, 1939). The passage to the omasum (and thence to the abomasum) is probably effected by the contractions of the reticulum as these are inhibited when the abomasum is distended.

The passage of ingesta through the *abomasum* and *small intestine* appears to be rapid. Czepa & Stigler (1929) found the emptying time of the abomasum of young sucking goats fed with milk and barium to be not greater than 3 hr., and although this does not necessarily represent the time spent by ingesta in the abomasum of an adult there is no obvious reason why the passage of ingesta through the abomasum should be delayed.

The secondary delay that occurs in the *large intestine* is not so great as that occurring in the rumen. Thus Lenkeit (1932) showed that the passage of stained particles of straw through lambs varied with the method of feeding; if the particles were given with oats, so that they passed to the rumen, maximal elimination occurred in the faeces from 48 to 72 hr. after feeding and persisted for 12 days or more. But if they were fed with milk, so that the food passed directly to the abomasum, then elimination was maximal within 12 hr. and persisted for only 5 days. The time spent in the rumen not only delayed the peak of elimination but caused the passage of the particles to be spread over a longer period. A comparison of the rate of passage of stained particles of oats or straw through the farm animals (in Table 1) indicates clearly the superiority of the stomach of the ruminant over the large intestine of the pig and horse as an organ for delaying the passage of food.

Table 1

	1st appearance	Maximal elimination	End of elimination
Horse (Lenkeit, 1933)	21–24 hr.	24 hr.	4–5 days
Pig (Lenkeit, 1933)	11–15 hr.	24 hr .	4–5 days
Sheep (Lenkeit, 1933)	14–19 hr.	48 hr. (still high at 72 hr.)	12 days (traces for days later)
Cow (Usuelli, 1933)	—	2–3 days	12–13 days

Spörri & Asher (1940) introduced barium into the caecum of a goat through a permanent fistula and found that the large gut cleared itself of barium within 8 hr. This rapid passage does not agree with the data on the passage of stained particles which must have been retained by the large gut for longer periods, as the passage through the abomasum and small intestine is relatively rapid. Normally when the caecum of a lamb is loaded with barium it lies upon the floor of the abdominal cavity, but this is not seen when only small quantities are present and the apex may then be high up in the flank owing to the presence of gas. If the caecum is fixed in the high position by a permanent cannula as in Spörri & Asher's animal, and the organ is then loaded with barium so that it would normally sink to the floor of the abdomen, the consequences are likely to be abnormal.

The ruminant can digest a greater proportion of the fibre of foodstuffs such as hay than animals relying on their caecum (Mangold, 1937; Ritzman & Benedict, 1938), and it is usual to attribute this to rumination, yet it is doubtful if this is true. Trautmann & Schmitt (1932) resected the reticulum from a kid goat and by so doing prevented the animal from ruminating, and although the goat was kept for 7 months on a normal ration with roughage it remained in good health although never seen to ruminate. The time per day devoted to rumination depends on the quantity and physical condition of roughage in the diet; thus animals fed entirely upon concentrated foodstuffs cease to ruminate (Ritzman & Benedict, 1938), while grinding hay into a fine consistency reduces the time devoted to rumination (Kick, Gerlaugh, Schalk & Silver, 1937). The reduction in size of food particles which is the object of remastication also has the effect of speeding the passage of food through the rumen; consequently the time during which fermentation of cellulose can occur is similarly reduced, and for this reason the act of rumination may actually reduce the quantity of cellulose digested. Kellner (1926) could find no difference in the digestibility of chopped or ground wheat and barley straw given to cattle, while recently Scharrer & Nebelsiek (1938) found that grinding soya-bean straw reduced the digestibility of the crude fibre for sheep.

The outstanding difference between the ruminant and mammals whose food ferments in the large gut is the time spent by the food in the alimentary tract, and the superiority of the rumen as an organ for delaying the passage of ingesta appears to be the most important factor in accounting for the greater digestion of fibre by these animals.

The physical conditions of the rumen. The conditions prevailing in the rumen are ideal for fermentation; the temperature remains constant within narrow limits and although a slight rise occurs after feeding and a slight fall during fasting, the variations do not exceed the extremes of 39·3 and 40·7° C. except when the animal drinks water (Krzywanek, 1929; Lampe, 1931). Fermentation in the rumen is anaerobic since oxygen is never found in more than small quantities. Its presence, together with that of sometimes quite considerable quantities of nitrogen, is attributed to air swallowed. The gases regularly present in quantity are carbon dioxide and methane. Of other gases hydrogen is often found and hydrogen sulphide and carbon monoxide may be present in traces (Washburn & Brody, 1937; Lugg, 1938; Dougherty, 1940, 1941).

The reaction of the ingesta of the rumen shows little variation in spite of the large quantities of organic acids produced. Electrometric measurements of the *p*H of samples of rumen contents have been made *in vitro* at regular intervals after feeding (Monroe & Perkins, 1939; Hale, Duncan & Huffman, 1940; Phillipson, 1942). Smith (1941) made *in vivo* measurements in cattle by means of a glass electrode. The reaction of the ingesta becomes more

acid after feeding, and this fall in *p*H is maintained for several hours, but later the *p*H rises and, if no further food is given, the rise is continued until presumably the *p*H of saliva is reached. The *p*H of the ingesta in both cattle and sheep is usually between 6 and 7, but higher figures are recorded, particularly in the early morning before feeding. Similarly, the *p*H may fall below 6 but not below 5 during the day, especially if much starchy or sugary food is given.

The stability of the *p*H of the ingesta is due to the saliva which is secreted in large quantities. The flow from the parotids is continuous in both the cow (Colin, 1886) and the sheep (Scheunert & Trautmann, 1921; Scheunert, Krzywanek & Zimmerman, 1930) and shows little response to feeding. The sublingual glands behave similarly, but the secretion of the submaxillary glands is variable and the response to mastication is copious. The total quantity of saliva produced by the cow in 24 hr. was estimated by Colin to be approximately 60 l. Similar measurements have not been made for small ruminants, but 5–42 c.c./10 min. is given for the quantity secreted from one parotid gland (Scheunert & Trautmann, 1921; Scheunert *et al.* 1930). The *p*H of saliva was investigated by Chrzaszcz & Schlechtlowna (1930) who found an average of 8·8 for the mixed saliva of cattle, while Weyers (1937) obtained 8·23–8·54 for the mixed saliva of sheep and estimated that the content of free and combined carbon dioxide was on an average 188 vol. %. The titratable value of saliva varies from 0·5 to 0·7 % calculated as sodium carbonate (Markoff, 1913; Scheunert & Krzywanek, 1930; Trautmann & Albrecht, 1931–2). The bases present in parotid saliva were sodium and potassium; carbonate, chloride and phosphate were also present in the ash, but calcium, magnesium and sulphate were absent. Scheunert & Krzywanek (1930) estimated that the sodium of the ash amounted to 30·25 %, potassium to 2·4 %, phosphate to 9·2 % and chloride to 1·9 %, while only traces of nitrogen were present. The saliva contains no diastatic enzyme (Markoff, 1913; Scheunert & Trautmann, 1921) and its functions are to provide (1) lubrication, for although the parotid and sublingual glands are serous the submaxillary glands are mixed and the secretion contains mucin (Scheunert & Trautmann, 1921); (2) sufficient fluid to keep the ingesta in the rumen in a semi-liquid condition, and (3) a buffer. The percentage dry matter obtained on desiccation of the rumen contents varies considerably, but the majority of figures recorded vary between 10 and 20 % for both sheep and cattle (Steinhauf, 1921; Nevens, 1928; Moon & Varley, 1942).

III. DIGESTIVE PROPERTIES OF THE RUMEN

Enzyme activity in the rumen. As might be expected, the breakdown and synthetic reactions demonstrated in this organ are numerous, although probably only a small proportion of those which are possible. Digestion in the rumen must be accomplished wholly by agents of external origin, since no secreting glands have been found and the saliva apparently contains no enzymes. Therefore the known agents at work are bacteria, protozoa and plant enzymes.

It is difficult to distinguish between plant enzymes, enzymes secreted by bacteria and enzymes liberated by the autolysis of dead bacteria or protozoa, any of which may be found free in the rumen and which may be responsible for such reactions as have been shown to take place in the presence of disinfectants.

The insoluble nature of cellulose, the chief carbohydrate constituent of the natural diet of the herbivore, suggests that the first stage in its breakdown is accomplished by an extracellular enzyme. Karrer & Staub (1924) showed that the fluid from cow's rumen, reticulum and omasum possessed the power of hydrolysing lichenin—a soluble polysaccharide of similar constitution to cellulose—to reducing sugar in the presence of toluol. Since they were able to show the presence of licheninase in various feeding stuffs including grains and grass, they believed that the enzyme activity in the cow's alimentary tract was of plant origin. The experiments of Karrer & Staub were presumably carried out at about the normal *p*H of rumen contents; at all events no statement is made to the contrary. Their demonstration of the activity of licheninase derived from plants was at *p*H 5·28; until it is shown that a plant licheninase can work at a higher *p*H, the assumption that such plant enzymes are active in the rumen must be received with caution. In any case the quantitative importance of such action has yet to be demonstrated. These authors found no licheninase activity in the abomasal contents of the cow, but some in the stomach contents of the pig. This latter result confirmed earlier work of Brown (1892) who demonstrated cellulase activity in the contents of the stomach of pigs and horses fed upon grain. Brown attributed the activity to an enzyme from the grain since he was able to exclude animal secretion and bacterial action as sources of the cellulase. This difference between activity in the bovine and the pig or horse may be due to the fact that in the abomasum, where fundic glands cover the whole body of the organ and the contents are fluid and well mixed, the enzyme may be destroyed by pepsin, while in the much more solid mass of recently ingested food in the cardiac end of the stomach of the pig or horse, which is devoid of fundic glands, no destruction of cellulase should take place and the enzyme may therefore have some activity.

It is more probable that the initial stages of cellulose breakdown in the rumen are brought about by enzymes of bacterial origin. Thus, hydrolysis of cellulose to glucose in the presence of a disinfectant by cultures from the rumen was shown by Woodman & Stewart (1928), and by cultures from mud by Pringsheim (1912). Pringsheim showed the same

effect by carefully raising the temperature of his culture, whereby his bacteria were killed and therefore fermentation checked, but the hydrolytic enzyme was not denatured. Pochon (1938) showed the formation of glucose from cellulose by anaerobes from mammalian intestine when the *p*H was reduced from 7·5 to 4·0. That hydrolysis to cellobiose and glucose is necessarily the first step in cellulose breakdown by bacteria has been doubted. Khouvine (1923) found that neither glucose nor cellobiose could replace cellulose as source of carbon for an anaerobe isolated from the human intestine. Cowles & Rettger (1931) studied a cellulose-splitting anaerobe isolated from various sources including cow's faeces, and found that while glucose was not attacked arabinose and xylose were not only attacked but were used in preference to cellulose. The suggestion has been advanced (Winogradsky, 1929; Boswell, 1941; Walker & Warren, 1938) on the basis of results obtained with aerobic organisms that oxidation of side-chain carbon atoms of the cellulose molecule giving a polyuronide type followed possibly by decarboxylation yielding pentosan precedes hydrolysis of the polysaccharide. Whether a similar process might take place anaerobically is not established; however, it would well explain the results of Khouvine and of Cowles & Rettger. The breakdown process of cellulose may well be different for different organisms and under different physical conditions; therefore the only reliable observations where rumen digestion is concerned are those made with rumen organisms under as nearly as possible rumen conditions.

Strong amylase activity in rumen contents has been shown by Sym, Stankiewicz & Zielinski (1939) who found that the enzyme responsible for the conversion of starch to dextrins and sugar in the presence of a disinfectant was associated with the solid portions of the ingesta, but could not be extracted by water. Sym believed the enzyme to be of plant origin and to be adsorbed on to the solid matter. Wegner, Booth, Bohstedt & Hart (1940) also report amylolytic activity of rumen ingesta.

The hydrolysis of xylan to xylose was shown by Iwata (1935) with mixed and pure cultures from the rumen. However, the extra-cellular nature of this change is not proved by his experiments, which were performed in the absence of a disinfectant. It is rather surprising under these circumstances that xylose rather than volatile acids was the end-product. However, the fact that a considerable amount of xylan remained even after 50 days' incubation suggests that the experimental conditions were unsatisfactory.

Lipase activity, again associated with the solid portions of the ingesta but not extractable by water, was shown by Sym *et al.* (1939) using olive oil as substrate. He believed it to be of plant origin. Sym (1938) has also shown the presence of a proteinase adhering to the solid residues, but extractable by water, in rumen contents. He suggested that the enzyme was of bacterial origin.

Fermentation of cellulose. The literature which is directly relevant to the problem of cellulose breakdown by bacteria in the herbivorous gut is scanty. Tappeiner (1884) inoculated cellulose suspended in 1 % meat extract with rumen contents. Gas production started on the first day, but the analyses of non-gaseous products were not made until 1–4 weeks after the fermentation was started. The products were then found to be carbon dioxide and methane with the former predominating; traces of formic acid and acetaldehyde, and large quantities of other volatile acids which, on the basis of the analysis of silver salts isolated from successive fractions of the steam distillate, he believed to be made up of at least 50 % of acetic acid and the rest butyric with a trace of propionic. From his figures it is possible to say that acids of higher equivalent weight than acetic are present in the first fractions of distillate and that some part of these at least must be butyric acid, but there is no real evidence to indicate that propionic acid is only present in traces. Tappeiner found, as subsequent workers have confirmed, that though alterations in media and physical conditions have little effect upon the non-gaseous products of cellulose fermentation, the composition of the gas is very sensitive to conditions. He found that the following conditions favoured the production of hydrogen and carbon dioxide instead of methane and carbon dioxide: substitution of meat extract by ammonium salts or amides as source of nitrogen, dilution and alkalinity. Tappeiner's use of untreated rumen contents as inoculum, which has not been followed by subsequent workers, probably gives a truer picture of the actual digestive process than results obtained with pure cultures; thus Khouvine (1923), in a detailed study of an organism from the human intestine, found that its power of fermenting cellulose was much increased by an unspecified contaminant. Woodman & Evans (1938) made repeated subcultures of rumen contents on shredded filter paper as source of carbon, the resulting cultures they used for studies of glucose and cellulose fermentations. The cellulose was found to give traces of pyruvic acid at an early stage and small quantities of lactic acid later; both tended to disappear towards the end of fermentation while volatile acids accumulated throughout. On the basis of the rate of steam distillation (Dyer, 1916–17) and unspecified chemical tests they state that the proportions of the acids vary considerably among formic, acetic and butyric, but that there is no evidence of the presence of propionic acid. Their results for glucose fermentation by the same cultures were qualitatively similar, though lactic acid accumulated in greater amount than was the case in cellulose fermentations. On the basis of these results and those already quoted (Woodman & Stewart, 1928), in which glucose was shown to be a product of cellulose breakdown in the

presence of toluol, these authors believe that glucose is a normal intermediate in cellulose fermentation and that pyruvic and lactic acids are also intermediates coming after glucose in the chain of reactions, which terminates with volatile acids in unknown proportions. Cultures of anaerobic organisms from allied sources, as for instance mud, sewage and river water, have been shown by various authors to ferment cellulose with the production of volatile acids in varying proportions, carbon dioxide, methane and sometimes hydrogen. It will be seen from the above survey that knowledge of the important subject of cellulose breakdown in the rumen is very scanty and further research is urgently required.

Fermentation of other carbohydrates. Even less is known about the breakdown of non-cellulose polysaccharides in the rumen. There is some evidence that anaerobic breakdown of 'hemicelluloses' in straw is more rapid than that of cellulose. Thus Acharya (1935) showed that pentosan and polyuronide constituents disappear more quickly than cellulose from straw soaked in water under the influence of bacteria which had accumulated spontaneously. Norman (1937) has suggested that in rotting plant material these carbohydrates form a readily accessible source of energy for the development of a culture of bacteria which then attacks the cellulose. Isolated straw hemicellulose is rapidly and completely broken down beyond the sugar stage by rumen bacteria (McAnally, 1942); however, its importance as a ready source of energy for bacteria is probably not so great in the rumen where a lively cellulose fermentation is proceeding continuously. The products of fermentation of these polysaccharides are probably similar to those formed from cellulose. Iwata's results have already been referred to. Seillière (1910) obtained eight parts of acetic to one part of butyric acid from xylan by the action of an inoculum from the large intestine of the guinea-pig. The fermentation of starch in the rumen has not been studied in detail. Volatile acids accumulated in the rumen contents of a sheep which had received a dose of 100 g. of starch (Phillipson & McAnally, 1942).

The fermentation of glucose in the rumen is important both on account of the presence of sugar in some feeding stuffs and also because glucose may well be an intermediate in the breakdown of polysaccharides. Woodman & Evans (1938) showed that glucose is broken down by rumen bacteria to pyruvic and lactic acids which then themselves disappear while volatile fatty acids accumulate. Phillipson (1942) found that lactic acid was produced *in vivo* when mangolds and cabbage were given as food and Phillipson & McAnally (1942) obtained lactic acid when glucose, fructose or sucrose but not maltose, lactose or galactose were introduced into the rumen through a fistula; they also confirmed the destruction of lactate *in vitro* and *in vivo*. The latter three sugars were fermented more slowly than the former, and it was suggested that in the former case the formation of lactic acid proceeds more rapidly than its breakdown, while in the latter case, since the sugars are not so rapidly attacked, lactic may be also formed as an intermediate but broken down as soon as it is produced.

The chemical changes which lead to the production of gases in the rumen are but little understood. These gases are carbon dioxide, methane, and some nitrogen which is presumably derived from swallowed air. Hydrogen, which is sometimes produced *in vitro*, is only found in traces *in vivo*. The production of methane was thought by early workers to be a characteristic of cellulose fermentation; however, methane is now believed to be produced by the reduction of carbon dioxide at the expense of a variety of organic substances acting as hydrogen donors (see Stephenson, 1939). While the carbon of methane is probably wasted so far as the animal is concerned, some of the carbon of carbon dioxide may yet be assimilated by bacteria and so find its way to the host animal in some form which may be of nutritive value. The fixation of nitrogen by rumen bacteria is unlikely unless the supply of nitrogen in other forms is low.

Synthesis in the rumen. Besides their power of breaking down complex substances to products which may be assimilated by the animal, the rumen bacteria and Protozoa synthesize material which may be of nutritive value to the host. Thus Baker & Martin (1938) have shown that the formation within the cell of a substance which stains with iodine in a similar manner to starch is characteristic of the cellulose-decomposing bacteria of the rumen, and of the caecum of other herbivores. The substance has not been chemically examined, but Baker believes that it is similar to starch; thus the bacteria would convert cellulose into a substrate for the intestinal amylases and make the carbon of cellulose available to the animal in the form of glucose. The nutritional importance of such a process would depend very much upon the quantitative aspect; that is, the proportion of cellulose carbon converted to the polysaccharide compared with that which reaches the animal in the form of volatile acids or other products of fermentation. This quantitative aspect has not so far been investigated.

It has not been possible to demonstrate satisfactorily that fat synthesis occurs in the rumen (Kraus, 1927; Krzywanek & Quittek, 1936). The literature concerned with the synthesis of protein from non-protein nitrogen by rumen bacteria is now extensive owing to the possible practical application of the process to animal feeding. It seems clear that two processes are involved: (1) if urea is the form of non-protein nitrogen used, deamination with the production of ammonia takes place rapidly (Lenkeit & Becker, 1938), (2) ammonia nitrogen is converted to protein by rumen bacteria (Wegner,

Booth, Bohstedt & Hart, 1940; Owen, Smith & Wright, 1941). This subject has recently been thoroughly reviewed by Goss (1943) who also deals fully with the subject of the synthesis of vitamins in the rumen, showing that the members of the B complex and also vitamin K are synthesized in the rumen.

Significance of the protozoa. None of the protozoa of the rumen have been extensively studied with a view to finding out their biochemical powers. Since a large number of nutritional types of protozoa exist (Doyle, 1943), it is not possible to make generalizations which will apply to all protozoa in the rumen; however, anaerobiosis and the absence of chlorophyll must be common to all. Kirby (1941) has reviewed the subject of the protozoa of the rumen and their activities. Their power of ingesting and breaking down starch has often been observed and particles of cellulose have been seen to disappear within the protozoan body, but the question of whether these protozoa can themselves digest cellulose has never been settled. In either case the argument has been advanced that digestion is accomplished by bacteria ingested with the carbohydrate (Ullman, 1932; Trier, 1926). Within a few hours of the ingestion of starch grains by protozoa of the rumen, the starch is digested and a glycogen-like substance is stored by the protozoon. This process seems to indicate breakdown by an amylase secreted by the protozoon followed by resynthesis of a polysaccharide rather than fermentation by bacteria whose products would less readily be used for synthesis of the glycogen-like body. This substance gradually disappears, being presumably used in providing energy for the protozoon; the chemical nature of the products of this reaction are unknown. Since much of the starch of a feed or dose is taken up by the protozoa (Usuelli, 1930), the manner in which they digest it is of importance in a consideration of the digestion of starch by the ruminant. Even if they were able themselves to digest cellulose the protozoa would not be of great importance in ruminant cellulose digestion since considerable bacterial action upon cellulose must take place before the plant fragments are reduced to a small enough size to be ingested by the protozoa.

Besides their power of feeding upon insoluble particles many protozoa subsist partially or wholly upon substances in solution. The protozoa of the rumen have been said to assimilate glucose and lactose (Trier, 1926; Westphal, 1934). It seems probable, however, that in the normal course of events any sugar in the feeding stuffs would be fermented by bacteria too rapidly for the amount stored by protozoa to be at all significant. The fact that for many protozoa—the 'Azetatorganismen' of Pringsheim (1937)—acetate can act either as sole source of carbon or a very favourable source is of considerable interest. Acetate is formed in large quantities by fermentation in the rumen, and if the protozoa should be capable of building it up into carbohydrate or fat within their bodies the host animal might thus be provided with a means of absorbing some part at least of the carbon of cellulose in the form of sugar or fat.

IV. COURSE OF DIGESTION IN THE RUMEN

Digestibility of food constituents. In the analytical procedure for the separation of crude fibre some cellulose is probably broken down and a small amount of other substances remains behind. However, the crude fibre fraction in digestibility trials may be taken to represent cellulose with a reasonable degree of accuracy. The digestibility of isolated cellulose is probably complete, since McAnally (1942) found that filter paper enclosed in silk bags and suspended in the rumen of a sheep disappeared in 4 days, while numerous trials in which partially isolated cellulose, in the form of straw treated with dilute alkali, has been given as food show digestibilities of 70–85 % of the crude fibre (Ferguson, 1942; Voltz, 1920; Fingerling & Schmidt, 1919).

The digestibility of cellulose as it occurs naturally is not always so great. Thus the digestibility of crude fibre in young grass is 84 % and in sugar-beet pulp is 90 %, while the crude fibre in grass at the hay stage is only 52 % digestible and in wheat straw 50 % (Woodman, 1930). This lowered digestibility appears to be accompanied by a greater content of lignin, and the usual assumption is made that the less digestible portion of the cellulose is in some way combined with the lignin. The process of treatment of straw with dilute alkali, which greatly increases its digestibility, is supposed to break this combination. Apart from the effect of its condition in the plant, cellulose may be rendered more or less digestible by the nature and amount of the other constituents of the food. Thus the majority of experimenters have found that both molasses and starch in a ration depress the digestibility of crude fibre (Mitchell & Hamilton, 1940; Zuntz, Heide & Klein, 1913; Armsby & Fries, 1918; Mertins, 1933). Mitchell & Hamilton (1940) quote results which indicate that crude fibre is better digested in a ration of low protein content than in rations containing more protein. In a synthetic diet containing a very little protein, however, the digestibility of cellulose was very poor (17·8 %), suggesting that a certain amount of nitrogen is necessary for the growth of the cellulose-splitting bacteria; this nitrogen could be supplied in the form of urea (Harris & Mitchell, 1941).

The digestibility of starch is high, thus Voltz (1916–17) found the digestibility coefficient of the nitrogen-free extractives of potato, which is largely starch, to be 85–95 % in the sheep. Iwata (1935) has collected references to digestibility trials upon pentosans. These results indicate that, as with cellulose, the more lignified materials give lower digestibility coefficients. Digestibility trials upon protein do not give a true picture of the extent to which food protein is broken down in the animal,

since some of the products of breakdown of plant protein may reappear as bacterial protein in the faeces. Calculations of the digestibility of crude protein in various foods of high protein content (based on figures in *Rations for Livestock*, 1939) indicate that 75–85 % disappears during passage through the animal. Similar calculations show that the oil fraction in feeding stuffs, in which the concentration of this fraction is high, is about 90 % digestible.

The rate and degree of digestion. The problem of estimating quantitatively the course of digestion in the rumen is not easy, owing to the continuous inflow of saliva and the continuous outflow of ingesta to the omasum, while in addition the products of digestion are continuously being absorbed. The use of non-toxic insoluble compounds mixed with the food as a marker with which the constituents of the food can be compared has only a limited value in ruminants, as the speed with which fine particles such as iron oxide or chromium trioxide leave the rumen does not necessarily correspond with the rate of passage of the individual constituents of the ration; indeed, Paloheimo (1939), who investigated the use of chromium trioxide, concluded that it was not possible to determine the time spent by a whole meal of concentrates in the rumen but only of particular physical fractions of the food. This difficulty can largely be overcome by using an indigestible fraction of the food itself as a marker provided that this is used to compare with other fractions of a single foodstuff such as hay, as the same criticism would apply if it were used as a marker in a mixed meal. Rathnow (1938) gave hay alone to sheep and used the iron content of the food as an indicator. This method proved satisfactory for, as far as could be judged, no iron passed from the food into solution in the rumen, the iron content of the liquid of the rumen remaining constant. The results indicated that 32 % of the total dry matter of hay was digested in the rumen within 3 hr., 48 % within 6 hr. and 59 % within 9 hr. Hale *et al.* (1940) employed the lignin contained in hay as an index with which to compare the other constituents; this should be a reliable indicator provided it is equally distributed throughout the plant and is completely indigestible. The results obtained on ingesta taken from the rumen of a cow with a fistula 14 hr. after feeding indicate that on an average 90 % of the total digestion of the dry matter of hay and 85 % of the total digestion of cellulose in hay occurs in the rumen while nitrogen-free extract disappears rapidly from the ingesta, in most cases up to 100 %, in 14 hr. The quantity of hay eaten had some influence upon the proportion of cellulose disappearing in the rumen as this was greater when 30 lb. of hay than when 20 or 10 lb. were given. The fact that 23 % of the lignin was lost during the passage through the whole alimentary tract throws some doubt upon the accuracy of these calculations, although arguments are put forward claiming that the loss must have occurred after the food left the rumen. Even if 10 % of the lignin is decomposed in the rumen, and there is evidence that bacteria of the rumen can attack this substance (McAnally, 1942), the loss of dry matter and cellulose from the rumen still remains considerable after allowing for this error. Protein also disappears up to 80–90 % in the rumen according to the lignin/protein ratio of the rumen ingesta and the total apparent digestion determined by digestibility trial.

Quantitative evidence regarding the methane evolved during fermentation is given by Washburn & Brody (1937), who found that the amount of methane excreted through the mouth and lungs by a dry cow on a maintenance ration of hay dropped from 2 l./kg. dry matter of the food during the first half-hour after feeding to 0·2 l./kg. 16–20 hr. after feeding. The wastage of digestible energy through the evolution of methane is variable and must be directly measured for correct energy balances to be struck, yet approximate figures are convenient for many calculations. Ritzman & Benedict (1938), after conducting fifty-two respiration experiments upon cows on a variety of diets, found that on an average 6·2 % of the total ingested energy or 9·6 % of the digestible energy was wasted as methane. Bratzler & Forbes (1940) state that the most significant relation between methane and the ration is the ratio of methane produced to carbohydrate digested, but the relationship is not a direct proportion as the amount of methane evolved is reduced relatively when the carbohydrate intake is increased. From their chart approximately 50 g. of methane are produced when 1000 g. of carbohydrate are digested, while approximately 250 g. of methane are produced when 6000 g. of carbohydrate are digested. The rapid decline in methane excretion found by Washburn & Brody indicates that most of the fermentation takes place within the first 24 hr. after feeding; during fasting, however, excretion of methane persists for 4 days when the previous food is hay (Ritzman & Benedict, 1938), and so presumably fermentation must persist for this time in the alimentary tract.

Indirect methods of a qualitative nature show that the peak of fermentation in the rumen is reached comparatively rapidly; thus Phillipson (1942) found that in sheep the concentration of total volatile acids in the ingesta of the rumen rose rapidly with a diet rich in starch or containing sugar in which cases the maximum concentration occurred 3–8 hr. later. When hay or grass, however, formed the diet, fermentation was slow and the maximum concentration did not occur until 12 hr. or more after feeding. The fluctuation in volatile acids was reflected by changes in *p*H. The volatile acids were maintained at a higher level and the *p*H at a lower level when sheep were grazing freely at pasture than when they were on a dry diet. This observation is in agreement with those of Smith (1941), who found a lower *p*H of the ingesta of the rumen in cattle when the animals ate grass than when they were stall fed.

The speed with which carbohydrates are fermented in the rumen varies; simple sugars such as

glucose are rapidly destroyed, and after the introduction of 100 g. of glucose into the rumen of sheep none is left after 2½ hr. (Phillipson & McAnally, 1942). Cane sugar produces similar results to glucose as regards acid production, but starch and cellulose produce only a slight rise in volatile acids although this is prolonged. The findings of Hale *et al.* (1940) that 85 % of the digestible cellulose disappears from the rumen within 14 hr. and those of Ritzman & Benedict (1938) that excretion of methane does not cease until 4 days after the last meal are apparently contradictory. The work of McAnally (1942), however, provides an explanation, as it shows that up to 40 % of the fibre of the faeces of sheep fed upon straw is potentially digestible, and that, of the potentially digestible cellulose of straw, only 25 % is digested as rapidly as in isolated cellulose while the remainder is less readily available. These observations applied to the normal and fasting animals suggest that the normal animal does not digest much of the less readily available material, since the ingesta do not stay long enough in the alimentary tract, but in the fasting animal the stagnation of food which occurs allows sufficient time for the bacteria to attack this less digestible material and so methane excretion is prolonged. If it is remembered that Hale's results refer to the material actually digested and not to that potentially digestible, then there is no contradiction between these data.

Absorption from the stomach. The first three compartments of the stomach are lined with stratified squamous epithelium, and this fact until recently has been sufficient to deter curiosity regarding absorption from these cavities, and it is usually stated categorically that no absorption occurs from any part of the stomach of the ruminant. This conception was somewhat shaken by the work of Trautmann (1933*a*), who showed that pilocarpine and atropine were rapidly absorbed through the epithelium lining the rumen. Rankin (1941), in addition, showed that several other drugs, and also glucose, could be absorbed from the rumen. The osmotic pressure of the fluid in the various parts of the stomach was found to be equal to that of the blood by Davey (1936), who put this forward as evidence in favour of absorption. A study of the formation and fate of the volatile acids in the rumen again leads to the conclusion that absorption of these substances must occur from the rumen and/or the omasum (Phillipson & McAnally, 1942), for although these acids are present in large quantities in the rumen little if any is present in the abomasal ingesta, and even the introduction of 50 g. of acetate or butyrate fails to increase significantly the concentration of these acids in the abomasum although the increased concentration in the rumen soon falls to the pre-dosing level. As these acids appear to be stable in the rumen ingesta, the only possible conclusion is that of direct absorption. Further investigation by radiological means (McAnally & Phillipson, 1942*a*) showed that a large molecule, such as that of sodium ortho-iodo-hippurate, could be absorbed from the rumen while the direct absorption of volatile acids from the rumen was proved by demonstrating the presence of a higher concentration of these acids in the blood draining the rumen than in the peripheral blood (McAnally & Phillipson, 1942*b*). Blood from the reticulum contained volatile acid in similar concentration to that found in blood draining the rumen, while blood draining the omasum contained somewhat less. Only traces were present in blood draining the abomasum and small intestine, but blood draining the caecum again carried a high concentration. This result, together with further unpublished work, shows that the concentration of volatile acids in the caecal ingesta is high compared with that in the contents of the abomasum or small intestine where only traces of the acids are found, and confirms the previous observation that volatile acids have largely disappeared from the ingesta by the time the material reaches the abomasum and that they reappear in the large gut where fermentation again becomes active. Since sugars and other readily fermentable dietary constituents will most probably have disappeared in the rumen, and starch, if any, in the rumen and small intestine, the source of the volatile acid in the caecum would seem to be cellulose. These results therefore appear to be in direct contrast with those of Trautmann & Asher (1939, 1941), who, as a result of experiments in which the rate of fermentation of cellulose was studied in the caecum of a goat with a caecal fistula, believe that cellulose does not remain long enough in the caecum for fermentation to set in. Trautmann and Asher's conclusions seem open to criticism on the following grounds: (1) as has already been pointed out, the time which a barium meal spends in a caecum fixed by a fistula high up in the flank and an ordinary meal in a normal caecum may be very different; (2) the plant material which these authors suspended in a perforated capsule in the caecum of their goat had not already been subjected to rumen digestion as cellulose in an ordinary meal would have been. It is possible therefore that normally cellulose digestion would set in more rapidly than in the capsule and that thus some considerable fermentation might occur within the estimated period during which the ingesta remain in the caecum. Trautmann & Asher themselves concede that this factor may be of some minor importance.

The quantity of volatile acid absorbed from the rumen and reticulum was estimated by McAnally & Phillipson (1942*b*) to vary between 2 and 4 g./hr. (expressed as acetic acid), but this figure must be regarded as minimal. A comparison of distillation curves of volatile acids from the ingesta of the rumen and from the blood draining the rumen indicates that a greater proportion of acetic acid to higher acids exists in the blood than in the ingesta. Experiments in which either acetate, propionate or butyrate in solution was introduced into the empty

rumens of sheep indicated that the rate of absorption appeared to depend upon the size of the molecule; thus after a standard time a higher concentration of acetate was found in the blood draining the rumen than of the other two acids, butyric acid being absorbed only in traces. Elsden (personal communication), using a chromatographic method of separating the acids, found that in the rumen 60 % of the total volatile acid is acetic, the remainder consisting of propionic, butyric and possibly a fourth as yet unidentified acid. In the blood draining the rumen, however, acetic amounts to 80 % of the total volatile acids, the remainder consisting of propionic and butyric acids.

It is not yet possible to balance the carbohydrate digested by the sheep against the products of digestion which, so far as is known, are mainly methane, carbon dioxide and a mixture of the lower fatty acids. Whether starch is largely converted to these substances, or whether much of it is digested by amylases in the small intestine, is not known. It is possible that a significant proportion is taken up by the ciliates of the rumen and so preserved from fermentation. Again, the significance of the bacterial polysaccharide described by Baker & Martin (1938) is not known as regards the total energy of the carbohydrate transported to the animal by this means. The lower fatty acids are produced in sufficient quantity and absorbed sufficiently rapidly to account for a part of the carbohydrate ingested. Their metabolism, therefore, about which but little is known, must be a subject of the first importance to those interested in ruminant digestion and nutrition.

V. SUMMARY

1. The diet of ruminants is peculiar in that it contains large quantities of cellulose and other carbohydrates which are indigestible except by the aid of bacteria; it is probable that the numerous protozoa of the rumen do not play a vital part in digestion.

2. Food passes to the fermentation chamber formed by the reticulum and rumen where it is kept in continual motion by rhythmic contractions of this organ. Regurgitation is caused by an inspiratory effort with a closed glottis; the bolus after being remasticated returns to the fermentation chamber.

3. Passage of ingesta from the rumen is slow and spread out over a period of days. Digestion is more efficient in the ruminant than in the horse and pig owing to the greater time spent by the food in its alimentary tract rather than owing to the act of remastication.

4. The voluminous secretion of saliva is responsible for the maintenance of a moisture content and reaction of the ingesta of the rumen suitable for the bacterial fermentation of carbohydrate. The conditions are anaerobic, since carbon dioxide and methane are the chief gases present, oxygen being found only in traces, if at all.

5. The breakdown of cellulose and other carbohydrates is accomplished mainly by bacteria, though plant enzymes may play a minor part. The products of this fermentation are largely carbon dioxide, methane and the volatile fatty acids, though it has been suggested that a starch-like polysaccharide is synthesized. The degree of digestion of cellulose is affected by the other constituents of the diet.

6. The rate and degree of digestion observed by the use of an indigestible constituent of the food as an index show that the total digestion and in particular that of cellulose in the rumen is considerable within 14 hr.; this is supported by measurements of the rate of evolution of methane.

7. Absorption of the volatile fatty acids and of certain drugs has been proved to occur from the rumen.

8. Fermentation is not confined to the rumen; it also occurs in the caecum of the ruminant.

The authors are greatly indebted to Sir Joseph Barcroft, F.R.S., for his constructive criticism and help in the preparation of the manuscript.

VI. REFERENCES

ACHARYA, C. N. (1935): *Biochem. J.* **29**, 528.
ARMSBY, H. P. & FRIES, J. A. (1918): *J. Agric. Res.* **15**, 269.
BAKER, F. & MARTIN, R. (1938): *Nature, Lond.*, **141**, 877.
BECKER, E. R. & EVERETT, R. C. (1930): *Amer. J. Hyg.* **11**, 362.
BECKER, E. R. & HSUING, T. S. (1929): *Proc. Nat. Acad. Sci., Wash.*, **15**, 684.
BECKER, E. R., SCHULZ, J. A. & EMMERSON, M. A. (1929–30): *Iowa St Coll. J. Sci.* **4**, 215.
BERGMAN, H. D. & DUKES, H. H. (1926): *J. Amer. Vet. Med. Ass.* **69**, 600.
BOSWELL, J. G. (1941): *New Phytol.* **40**, 20.
BRATZLER, J. W. & FORBES, E. B. (1940): *J. Nutrit.* **19**, 611.
BROWN, H. T. J. (1892): *Chem. Soc.* **61**, 352.
CHRZASZCZ, T. & SCHLECHTLOWNA, Z. (1930): *Biochem. Z.* **219**, 30.
COHENDY, M. & WOLLMAN, E. (1914): *C.R. Acad. Sci., Paris*, **158**, 1283.
COLIN, G. (1886): *Traité de Physiologie Comparée des Animaux*, 3rd ed. vol. **1**. Paris.
COWLES, P. B. & RETTGER, L. F. (1931): *J. Bact.* **21**, 167.
CZEPA, A. & STIGLER, R. (1926): *Pflüg. Arch. ges. Physiol.* **212**, 300. — (1929): *Fortschr. naturw. Forsch.* **6**, 1.
DAVEY, D. G. (1936): *J. Agric. Sci.* **26**, 328.
DOUGHERTY, R. W. (1940): *J. Amer. Vet. Med. Ass.* **96**, 43. — (1941): *J. Amer. Vet. Med. Ass.* **99**, 110.
DOYLE, W. L. (1943): *Biol. Rev.* **18**, 119.
DYER, D. C. (1916–17): *J. Biol. Chem.* **28**, 445.
FANTHAM, H. B. (1922): *S. Afr. J. Sci.* **19**, 332.
FERGUSON, W. S. (1942): *Biochem. J.* **36**, 786.
FINGERLING, G. & SCHMIDT, K. (1919): *Landw. VersSta.* **94**, 114.
FLOURENS, P. (1844): *Mémoires d'Anat. et de Physiol. Comp.* Paris.
GLIMSTEDT, G. (1936): *Skand. Arch. Physiol.* **73**, 48.
GOSS, H. (1943): *Nutr. Abstr. Rev.* **12**, 531.
HALE, E. B., DUNCAN, C. W. & HUFFMAN, C. F. (1940): *J. Dairy Sci.* **23**, 953.
HARRIS, L. E. & MITCHELL, H. H. (1941): *J. Nutrit.* **22**, 167.
HAUBNER, G. C. (1837): *Über die Magenverdauung der Wiederkäuer*. Vienna.
IWATA, H. (1935): *Bull. Coll. Agric. For. Moriaka*, no. 21.
KARRER, P. & STAUB, M. (1924): *Helv. chim. Acta*, **7**, 916.

KELLNER, O. (1926): *The Scientific Feeding of Animals*, p. 103. (trans. W. Goodwin), 2nd ed. London.
KHOUVINE, Y. (1923): *Ann. Inst. Pasteur*, **37**, 711.
KICK, C. H., GERLAUGH, P., SCHALK, A. F. & SILVER, E. A. (1937): *J. Agric. Res.* **55**, 587.
KIRBY, H. (1941): *Protozoa in Biological Research*, p. 973. Columbia Univ. Press.
KRAUS, F. J. (1927): *Biol. gen.* **3**, 347.
KRZYWANEK, F. W. (1929): *Pflüg. Arch. ges. Physiol.* **222**, 89.
KRZYWANEK, F. W. & QUAST, P. (1937): *Pflüg. Arch. ges. Physiol.* **238**, 333.
KRZYWANEK, F. W. & QUITTEK, G. (1936): *Biederm. Zbl. (Tierernähr.)*, **8**, 136.
KUSTER (1914): *Arb. GesundhAmt.*, Berl., **48**, 1.
LAMPE, K. W. (1931): Inaug. Diss. Leipzig.
LENKEIT, W. (1932): *Berl. tierärztl. Wschr.* **48**, 18. — (1933): *Ergebn. Physiol.* **35**, 573.
LENKEIT, W. & BECKER, M. (1938): *Z. Tierernähr. Futtermittelk.* **1**, 97.
LUGG, J. W. H. (1938): *J. Agric. Sci.* **28**, 688.
McANALLY, R. A. (1942): *Biochem. J.* **36**, 392.
McANALLY, R. A. & PHILLIPSON, A. T. (1942*a*): *J. Physiol.* **101**, 13*P*. — (1942*b*): *J. Physiol.* **101**, 14*P*.
MAGEE, H. E. (1932): *J. Exp. Biol.* **9**, 409.
MANGOLD, E. (1937): *L'Utilisation Alimentaire de la Cellulose*. Paris: Hermann et Cie.
MANGOLD, E. & RADEFF, T. (1930–1): *Arch. Tierernähr. Tierz.* **4**, 173.
MARKOFF, J. (1913): *Biochem. Z.* **57**, 1.
MERTINS, H. (1933): *Z. Tierz. ZüchtBiol.* **26**, 367.
MITCHELL, H. H. & HAMILTON, T. S. (1940): *J. Agric. Res.* **61**, 847.
MONROE, C. F. & PERKINS, A. E. (1939): *J. Dairy Sci.* **22**, 983.
MOON, F. E. & VARLEY, N. (1942): *Vet. Rec.* **54**, 359.
NEVENS, W. B. (1928): *J. Agric. Res.* **36**, 777.
NORMAN, A. G. (1935): *J. Agric. Sci.* **25**, 529. — (1937): *The Biochemistry of Cellulose, the Polyuronides Lignin, etc.* Oxford.
NORMAN, A. G. & RICHARDSON, H. L. (1937): *Biochem. J.* **31**, 1556.
NUTTALL, G. H. F. & THIERFELDER, H. (1895–6): *Hoppe-Seyl. Z.* **21**, 109.
OWEN, E. C., SMITH, J. A. B. & WRIGHT, N. C. (1941): *Nature, Lond.*, **147**, 710.
PALOHEIMO, L. (1939): *Biederm. Zbl. (Tierernähr.)*, **11**, 370.
PASTEUR, L. (1885): *C.R. Acad. Sci., Paris*, **100**, 68.
PHILLIPS, M., DAVIS, B. L. & WEIHE, H. D. (1942): *J. Agric. Res.* **64**, 533.
PHILLIPSON, A. T. (1939): *Quart. J. Exp. Physiol.* **29**, 395. — (1942): *J. Exp. Biol.* **19**, 186.
PHILLIPSON, A. T. & McANALLY, R. A. (1942): *J. Exp. Biol.* **19**, 199.
POCHON, J. (1938): *Trav. Sta. zool. Wimereux*, **13**, 575.
POLYANSKY, G. & STRELKOW, A. (1935): *Trav. Inst. biol. Peterhof*, **13–14**, 68 (quoted from Kirby, H., listed above).
PRINGSHEIM, E. G. (1937): *Nature, Lond.*, **139**, 196.
PRINGSHEIM, H. (1912): *Hoppe-Seyl. Z.* **78**, 266.
QUIN, J. I. & VAN DER WATH, J. G. (1938): *Onderstepoort J. Vet. Sci.* **11**, 361.
QUIN, J. I., VAN DER WATH, J. G. & MYBURGH, S. (1938): *Onderstepoort J. Vet. Sci.* **11**, 341.
RANKIN, A. D. (1941): Thesis Cornell (*Abstr. Vet. Bull.* **11**, 328).
RATHNOW, H. D. (1938): Inaug. Diss. München.
RATIONS FOR LIVESTOCK (1939): *Bull. Minist. Agric.* no. 48.
RITZMAN, E. G. & BENEDICT, F. G. (1938): *Publ. Carneg. Instn*, no. 494.
SCHALK, A. F. & AMADON, R. S. (1928): *Bull. N. Dak. Agric. Exp. Sta.* no. 216.
SCHARRER, K. & NEBELSIEK, H. (1938): *Z. Tierernähr. Futtermittelk*, **1**, 72.
SCHEUNERT, A. & KRZYWANEK, F. W. (1930): *Pflüg. Arch. ges. Physiol.* **223**, 472.
SCHEUNERT, A., KRZYWANEK, F. W. & ZIMMERMAN, K. (1930): *Pflüg. Arch. ges. Physiol.* **223**, 462.
SCHEUNERT, A. & TRAUTMANN, A. (1921): *Pflüg. Arch. ges. Physiol.* **192**, 33.
SEILLIÈRE, G. (1910): *C.R. Soc. Biol., Paris*, **68**, 989.
SMITH, V. R. (1941): *J. Dairy Sci.* **24**, 659.
SPÖRRI, H. & ASHER, T. (1940): *Schweiz. Arch. Tierheilk.* **82**, 204, 243.
STEINHAUF, P. (1921): Inaug. Diss. Berlin.
STEPHENSON, M. (1939): *Bacterial Metabolism*, p. 59. Longmans.
STIGLER, R. (1931): *Arch. Tierernähr. Tierz.* **4**, 613.
SYM, E. A. (1938): *Acta Biol. exp., Varsovie*, **12**, 192.
SYM, E. A., STANKIEWICZ, W. & ZIELINSKI, F. (1939): *Enzymologia*, **6**, 113.
TAPPEINER, H. (1884): *Z. Biol.* **20**, 52.
TOUSSAINT, H. (1875): *Arch. Physiol. Norm. Path.* **2** (2nd series), 141.
TRAUTMANN, A. (1933*a*): *Arch. Tierernähr. Tierz.* **9**, 178. — (1933*b*): *Arch. Tierernähr. Tierz.* **9**, 575.
TRAUTMANN, A. & ALBRECHT, H. (1931–2): *Arch. wiss. prakt. Tierheilk.* **64**, 93.
TRAUTMANN, A. & ASHER, T. (1939): *Z. Tierernähr. Futtermittelk.* **3**, 45 (*Nutr. Abstr. Rev.* 1940–1, **10**, 343). — (1941): *Arch. wiss. prakt. Tierheilk.* **76**, 317.
TRAUTMANN, A. &. SCHMITT, J. (1932): *Arch. wiss. prakt. Tierheilk.* **65**, 559. — (1935): *Dtsch. tierärztl. Wschr.* **43**, 177.
TRIER, H. J. (1926): *Hoppe-Seyl. Z.* **4**, 305.
ULLMANN, T. (1932): *Hoppe-Seyl. Z.* **17**, 520.
USUELLI, F. (1930): *Clin. Vet. Milano*, **53**, 543, 625, 787. — (1933): *Profilassi*, **6**, 7.
VOLTZ, W. (1916–17): *Landw. Jb.* **50**, 455. — (1920): *Biochem. Z.* **102**, 151.
WALKER, E. & WARREN, F. L. (1938): *Biochem. J.* **32**, 31.
WASHBURN, L. E. & BRODY, S. (1937): *Bull. Mo. agric. Exp. Sta.* no. 263.
WATSON, R. H. (1941): *Aust. Vet. J.* **17**, 52.
WEGNER, M. I., BOOTH, A. N., BOHSTEDT, G. & HART, E. B. (1940): *J. Dairy Sci.* **23**, 1123.
WESTER, J. (1926): *Die Physiologie u. Pathologie der Vormagen beim Rinde*. Berlin: Julius Springer.
WESTPHAL, A. (1934): *Z. Parasitenk.* **7**, 71.
WEYERS, H. (1937): Inaug. Diss. Berlin.
WINOGRADOW, M., WINOGRADOWA-FEDOROWA, T. & WERENINOW, A. (1930): *Zbl. Bakt.* Abt. II, **81**, 230.
WINOGRADSKY, S. (1929): *Ann. Inst. Pasteur*, **43**, 549.
WISE, G. H. & ANDERSON, G. W. (1939): *J. Dairy Sci.* **22**, 697.
WOOD, T. B. & BERRY, R. A. (1905): *J. Agric. Sci.* **1**, 176.
WOODMAN, H. E. (1930): *Biol. Rev.* **5**, 273.
WOODMAN, H. E. & EVANS, R. E. (1938): *J. Agric. Sci.* **28**, 43.
WOODMAN, H. E. & STEWART, J. (1928): *J. Agric. Sci.* **18**, 713.
ZUNTZ, N., v. D. HEIDE, R. & KLEIN (1913): *Landw. VersSta.* **79–80**, 781.

THE INVESTIGATION OF PLANT NUTRITION BY ARTIFICIAL CULTURE METHODS

By D. I. ARNON and D. R. HOAGLAND,
Laboratory of Plant Nutrition, University of California

(*Received* 28 *October* 1943)

The environment of a higher plant growing in a soil is one of extraordinary complexity and the investigation of the nutrition of the plant in its normal habitat is correspondingly difficult. Ultimately the whole system of the soil, the plant and the atmosphere must be understood before many of the objectives of research in plant nutrition can be attained. It has long been recognized that basic scientific principles of plant growth are not possible of adequate development without experimentation in which control and simplification of certain environmental factors can be introduced. One step in this direction, of outstanding significance, consists in substituting an artificial and controllable root medium for the complex and heterogeneous natural medium of a soil. The purpose of the present review is to survey the techniques of artificial culture and to indicate a few of the advances in plant physiology that have depended on their utilization. Many species of higher plants have been grown experimentally in artificial culture with the roots furnished only with a solution of inorganic salts, under suitable conditions of root aeration. No factor inseparable from the soil, nor any preformed organic compound, externally derived, including vitamin substances, appears to be indispensable to the normal functions of the plants, although research has been chiefly directed to limited groups of plants. The importance of organic matter and of clay colloids in soil is of course beyond question, but these factors may be regarded as operating in a secondary sense and should be distinguished from those factors that are indispensable for plant growth in a primary sense.

I. HISTORICAL SKETCH

The history of growing plants without soil illustrates that even in science a technique may be used before its underlying principles are understood or even correctly surmised. The earliest scientific record of dispensing with the soil as a medium of plant growth dates back to the end of the seventeenth century. In 1699 Woodward reported in a notable paper a series of experiments with spearmint grown in rain, river and conduit water, to which in some cases garden mould was added. The purpose was to disprove the view based on an earlier experiment of van Helmont that water is the 'principle of vegetation', that is to say, the sole substance from which all the plant is derived. Woodward's experiments showed that plants grew best in those of his water cultures which contained the greatest admixture of 'terrestrial matter'. His conclusion was that 'vegetables are not formed of water: but of certain terrestrial matter'. Thus the search for a spurious 'principle of vegetation' led to the inadvertent discovery, not fully appreciated until a century and a half later, that plants can be grown in an aqueous medium without soil. Another early instance is the work of Duhamel du Monceau in 1758, who grew bean, chestnut, almond, and oak plants in water.

The real development of the technique of water culture had to await the formulation of modern concepts of chemistry initiated by Lavoisier in the latter half of the eighteenth century and the subsequent advance in plant physiology. The discovery of oxygen by Priestley, who also pointed out the purifying effect of plants on air 'vitiated' by animals, and the overthrow of the phlogiston theory, coupled with the elucidation of the nature of combustion and of the composition of water and carbonic acid by Lavoisier, was promptly followed by the formulation of the concept of photosynthesis by Ingenhousz, and by Senebier. Thus for the first time in the long association of man with green plants the peculiar nature of their nutrition was revealed. Soon following were the striking experiments of Theodore de Saussure, the results of which, when published in 1804, provided definite evidence for the modern views on plant nutrition, namely, that plants are made of chemical elements derived from three sources, air, water, and soil, and that plants grow and increase in size and weight by synthesizing from these sources of nutrition various plant substances. Among the many contributions of de Saussure, the one particularly relevant to the present discussion was the demonstration that although the ash constituents represent a quantitatively small proportion of the weight of the plant, they are none the less essential for growth and development.

Despite the solid foundation on which the newly gained insight into the nature of plant nutrition rested, the four decades following the publication of de Saussure's work witnessed a relapse into the

vitalistic and 'humus' theories of plant physiology which envisaged the absorption by the plant of ready-made organic food from the soil and attributed no significance to the ash constituents. Illustrative of the prevailing uncertainty in thought was the offer in 1838 of a prize by the Academy of Göttingen for the best answer to the question: 'Are the so-called inorganic elements which are found in the ash of plants still to be found there when they are not supplied to the plant from without, and, whether these elements are such essential constituents of the vegetable organism as to be required for its full development?' It was not till the appearance in 1840 of Liebig's book on the relation of chemistry to agriculture and physiology that the significance of Ingenhousz's and de Saussure's work was forced on the attention of the plant physiologists and chemists of that period. Although Liebig failed to understand the role of soil as a source of nitrogen, he forcefully marshalled the evidence for the view that the soil is the source of the ash constituents for the plant, and that these represent elements indispensable for plant growth. Once this was recognized it was only natural to attempt to supply these elements and water independently of soil. The credit for initiating exact experimentation in this field belongs to Jean Boussingault, who is regarded as the founder of modern methods of conducting experiments on vegetation. Boussingault, who had begun his work even before 1840, grew plants in insoluble artificial soils: sand, quartz, and sugar charcoal, which he watered with solutions of known composition. His results provided experimental verification for the mineral theory of plant nutrition and were at once a demonstration of the feasibility of growing plants in a medium other than a 'natural soil'. In later years Boussingault also established the importance of nitrates as the chief source of nitrogen for higher plants. The method of growing plants in artificial insoluble soils was subsequently improved by Salm-Horstmar, and has frequently been used since.

After plants had been successfully grown in artificial culture media it was but one more step to dispense with any solid medium and attempt to grow plants in water to which the requisite chemical elements were added. This was accomplished by Sachs who published in 1860 a standard solution formula for growing plants in water culture, and, along with his contemporary Knop, who was simultaneously working on this problem, opened the modern period of investigation in plant nutrition. Among the many formulae for nutrient solutions used in the period following 1860 that proposed in 1865 by Knop became, with modifications, one of the most widely employed in studies of plant nutrition. The experimentation based on nutrient solution technique in the half-century following the contributions of Sachs and Knop, was exhaustively reviewed by Tottingham (1914). Livingston (1942) has compared the properties of different nutrient solutions.

II. METHODS

Water culture. Historically the most prominent, the water-culture method continues to maintain its usefulness for the study of a number of diverse problems in which control of the nutrient medium is essential. Its chief advantage lies in the elimination of solid matter around the roots and the consequent simplicity of experimental technique in maintaining a known concentration and composition of nutrient solution. The essentials of this technique have, with the important exceptions of aeration and the addition of micronutrients, remained the same as those developed by Sachs and Knop. The seed is germinated in sand, sawdust, or on cheesecloth, blotters, filter paper, and the like, in such a manner that the small seedlings can be easily transferred, without damage to the roots, to the nutrient solution. If the container used is capable of absorbing salts (crocks, wood), suitable safeguards (glazing, non-toxic coating) must be employed to prevent complications from deposition of salts from previous experiments. For the highest degree of precision pyrex glass is found to be satisfactory. Recent improvements in aeration technique have removed the disadvantage of deep containers even for plants possessing high oxygen requirements (Hoagland & Arnon, 1938; Colby, 1933). The absence of solid matter from the nutrient substrate deprives the roots of anchorage and makes it necessary to provide special means for supporting the plants in the culture vessel. Separate provisions must, of course, be made for plants like potatoes and sugar beets which are grown for their subterranean organs. These plants can be supported in a frame several inches deep to one end of which a wire screen is attached, the screen being placed directly above the level of the nutrient solution. Tubers develop in a bed of loose material in the frame. In a similar way water-culture technique can include a porous bed superimposed over the nutrient solution. This type has been designated by the term 'hydroponics' and has been applied to the cultivation of plants of wide variety by Gericke (1940), who first devised it. Some writers have adopted the term 'hydroponics' to refer to any kind of culture in which the roots of the plant are immersed in a liquid nutrient medium.

Until recently water culture, although used successfully for numerous species of plants by a host of workers, has sometimes given only indifferent results or even failure. In the light of present knowledge two principal causes of failure can be singled out. First, there was a possibility that the nutrient solution was either qualitatively or quantitatively deficient in one or more of the micronutrient elements (Brenchley, 1943), and secondly, that no adequate provision was made for aeration, although some earlier investigators considered this factor. The two most satisfactory methods for insuring an adequate supply of oxygen to the roots, are, in our experience,

sintered pyrex glass aerators (Furnstal & Johnson, 1936) and porous carbon tubes (Arnon & Hoagland, 1940). The sintered glass aerators can be prepared on a large scale with the aid of ordinary laboratory facilities (Furnstal & Johnson, 1936). Adequate aeration is often automatically brought about by certain modifications of the standard water-culture technique, such as drip or flowing culture (Shive & Stahl, 1927; Shive & Robbins, 1938). It is also possible to increase the supply of oxygen to the roots by lowering the level of the nutrient solution, but this alone may not insure an optimal supply of air for roots of plants with a high oxygen requirement, even if shallow containers with large exposed surfaces are used (Arnon & Hoagland, 1940). Nevertheless, the porous-bed technique of Gericke will often provide adequate aeration for crop production, comparable with that of a highly fertile soil.

It should be recalled that plants differ in their oxygen requirements, and certain species, such as rice and willow for example, may be injured by forced aeration of the nutrient solution which, however, is beneficial for the growth of other species such as the tomato. The quantitative oxygen requirements of three different species have recently been reported by Gilbert & Shive (1942).

Sand or gravel culture. The use of an inert, solid substrate in conjunction with a nutrient solution affords a root environment similar in some physical respects to soil, and is well adapted to certain experiments, and especially to the recently proposed commercial growing of plants without soil. The coarse nature of the medium automatically insures liberal aeration without the use of special devices and provides anchorage for the roots, thus dispensing with the need for support for the aerial portions of those plants which ordinarily do not require it when grown in soil. Among the considerations which limit the use of sand culture in some types of controlled experimentation are: the introduction of a solid phase, with associated problems of surface forces; an intermittent rather than a continuous supply of nutrients owing to repeated irrigation; a large proportion of container space occupied by inert material rather than by nutrient solution; uneven concentration of nutrients surrounding different portions of the root system due to more rapid absorption in regions of the densest root growth; inaccessibility of the roots to continuous inspection during growth; and lastly, the difficulty of complete removal of root debris and salt contaminants if the sand or gravel is to be put to repeated use. None of these limitations, however, is critical enough to prevent the successful application of sand culture in many investigations in plant nutrition.

Various inert materials have been found suitable; of which sand continues to be the most widely used. The kind of sand is important in many cases. For most accurate work, particularly with micronutrients, purity is essential and only pure quartz alone is suitable. In experiments on inorganic nutrients required by plants in large amounts, the chief factor in the quality of sand is usually particle size. With too fine particles drainage is slow, and with too coarse particles excessive irrigation may be necessary because of the small amount of water held against gravity. Least desirable is sand with particles of such mixed sizes that little pore space remains after the sand has settled. The most suitable sand and its preparation are discussed by Eaton (1941). In recent years considerable attention has been given to a number of inexpensive inert materials other than sand suited to artificial plant culture, particularly on a large scale. Materials such as gravel, cinders, granite chips, and a burnt clay and shale compound known as Haydite have been employed, often with excellent results. The use of each of these materials has been discussed by Withrow and others (1938, 1943) and by Laurie & Kiplinger (1940). The variety of containers for sand and gravel culture is large. For small installations, metal, glass and earthenware containers, usually designed to work with automatic delivery devices for the nutrient solution, are commonly selected (Shive & Robbins, 1938; Eaton, 1930, 1941–2; Chapman & Liebig, 1938). For larger installations waterproofed wooden benches (Withrow & Biebel, 1936) and concrete beds are employed (Laurie & Kiplinger, 1940; Eaton, 1930).

Methods for supplying nutrients. In ordinary water cultures the nutrient salts may be replenished from time to time as they are absorbed by the plants. With large plants, and particularly when climatic conditions favour transpiration, water is depleted much more rapidly than the dissolved salts. In certain critical experiments, which require the maintenance of a particular range of concentration for a given nutrient, the addition of salts to the nutrient solution is guided by periodic chemical analyses. If the plan of the experiment demands it, the constancy of composition of the nutrient solution during the experimental period can be maintained either by continuously flowing nutrient solutions (Shive & Stahl, 1927; Shive & Robbins, 1938; Trelease & Livingston, 1922; Johnston, 1927; Wean, 1935; Johnston & Hoagland, 1929; Zinzadze, 1935; Zurbicki, 1933) or by providing a large and continuously stirred volume of solution (Arnon, 1937; Parker & Pierre, 1928), so that changes in the nutrient solution induced by the absorption of ions by the plant are distributed over a large volume and rendered quantitatively unimportant. An adequately rapid rate of renewal is also necessary when flowing nutrient solutions are used. The simplest procedure for supplying nutrients to a sand culture consists of periodic manual application of a standard nutrient solution combined with occasional flushing with water to remove the accumulation of unabsorbed salts. However, in large experimental installations this time and labour-consuming method has been replaced by automatic devices requiring relatively little attention (Shive & Stahl, 1927; Zurbicki, 1933; Shive &

Robbins, 1938; Trelease & Thomson, 1935). With small culture vessels, drippers in large variety, often made of capillary tubing, have been adapted to convey the nutrient solution by siphoning. Recently more elaborate and fully automatic devices have been described for both large and small culture vessels (Eaton, 1936; Chapman & Liebig, 1938). In some, clock-controlled valves make possible the pumping of nutrient solution over the sand beds at selected time intervals.

The recent applications of sand-culture methods to commercial growing of greenhouse crops in gravel and similar inert materials often depends on an automatic subirrigation method for supplying the nutrients (Withrow & Biebel, 1936). The nutrient solution is supplied through a perforated pipe laid at the *bottom* of a watertight shallow bed, such as a greenhouse bench, which is filled with gravel. The delivery pipe is connected to a centrifugal pump which circulates the solution from a reservoir placed at a lower level than the bed. After each flooding the solution drains by gravity back into the reservoir. The pump is automatically operated at regular intervals by a time clock. Another application of the sand-culture system to commercial production was described by Biekart & Connors (1935). One point demanding attention in automatic installations is the selection of materials; pipes, tanks, paints, etc. free from ingredients toxic to plants. Galvanized iron should in general be avoided because of zinc toxicity. Black iron or cement-lined pipes have been found satisfactory. Black asphaltum paint (derived from a petroleum base), having been found physiologically inert, is widely used in plant culture work to coat exposed metal surfaces.

Composition of nutrient solutions. Since the utilization of inorganic nutrients, especially nitrogen, depends on the products of photosynthesis, it is clear that the suitability of a given solution is linked to climatic conditions, which influence photosynthesis and the general metabolic processes of the plant. Despite the early recognition by the originators of the water-culture technique (Sachs, 1887) that fairly wide latitude is permissible in the preparation of nutrient solutions, suggestions have been made that for a particular climate, if not for most climates, some physiological basis may be used as a guide in devising a special composition of a nutrient solution of unusual merit. Different concepts are sometimes advanced for aiming at an 'optimal' nutrient solution, among them that there should be a 'balance' between the various nutrients.

An analogy with soil conditions offers no compelling reason for either selecting or maintaining an originally selected 'best' nutrient solution. Plants in nature do not grow in nutrient media of fixed composition. Indeed the opposite is the case, for not only do plants grow in soils widely differing in their power of supplying nutrients, but also there occur in any one soil wide seasonal fluctuations in the nutrients that were available at the beginning of the growing season. It is also obvious that as different species of plants may grow well in one kind of soil, the same nutrient solution can support more than one kind of plant. Neither does the composition of a plant afford a universal guide in devising nutrient solutions. Plant composition is in part genetically determined and is influenced by climatic factors. Different kinds of plants grown at the same time in the same soil or nutrient solution will develop tissues with different proportions of inorganic constituents. Moreover, the inorganic composition of various parts of a plant may vary considerably and different nutrient solutions may be suggested by the composition of the fruit and of the foliage. Furthermore, it is evident that since the composition of the nutrient solution affects, within certain limits, the composition of the plant, particularly that of its vegetative organs, it is not logical to assume that an analysis of the plant affords an accurate basis for formulating a nutrient solution of unique physiological properties.

Considerable difficulty is also encountered in basing the composition of a nutrient solution on some originally 'balanced' proportion of the inorganic nutrients. The absence of physiological criteria capable of defining this balance with precision has already been mentioned. It is also true that, owing to the differential absorption of nutrients by plants, no one nutrient solution retains its identity after plants have grown in it for some time, unless very large volumes of solution relative to the numbers of plants are made available, as by flowing solutions. The experience of many workers has shown that different nutrient solutions are capable of supporting excellent plant growth. A crucial factor from the standpoint of the optimal development of the plant is the supply of a *sufficient quantity of each essential element* within suitable total concentration ranges and fairly broad limits of ionic proportions, granted that some cases may arise in which special attention must be given to the effect of one ion on the absorption of another. For example, the relation of iron concentration to manganese concentration has been stressed by some investigators (Shive, 1941). At this point the discussion has been mainly concerned with the growth and yield of plants, rather than with their composition, as affected by solution composition, although the latter question is of great interest in relation to the use of plants by animals.

An attempt was recently made (Arnon & Hoagland, 1940) to devise a 'good' nutrient solution, not from the standpoint of conferring unusual benefits on the plant, but rather from that of greatest convenience in management of artificial water cultures. The objective was to find a nutrient solution which would meet the following requirements: (*a*) in its proportion of nutrients the actual removal of ions by the plants should be approached, although, as already suggested, this is feasible to only a limited extent, considering the changing action of the plant at dif-

ferent physiological stages; (*b*) its *p*H should remain within a range favourable for plant growth and should require little attention between changes or replenishments of the nutrient solution. The composition of this solution was based on the amounts of the various nutrients absorbed by tomato plants, as determined by a series of periodic analyses of nutrient solution for a period of five months. The new solution, devised in this manner, has been found capable of supporting excellent growth of several kinds of plants with a minimum of attention and expense in replenishing the absorbed salts.

Supply versus concentration of nutrients. A consideration sometimes overlooked in comparing effects of different nutrient solutions is that of continuous supply of a nutrient in contrast to concentration at a given time. If other factors are alike, the higher the initial concentration of an element in a given volume of a nutrient solution, the larger will be the supplying power of the nutrient medium for this particular nutrient. When small volumes of solution are used over an extended time period, a low concentration of a rapidly absorbed element like potassium may lead to early depletion of the element concerned. One nutrient solution may be judged qualitatively inferior to another which has an initially higher concentration of potassium because of volume relationships. The importance of large volumes or continuous renewal of a nutrient solution is obvious in this connection (Johnston & Hoagland, 1929; Parker & Pierre, 1928).

A question is sometimes raised as to the possibility that the entire requirement of nutrient salts needed to carry plants to maturity might be initially incorporated in the nutrient solution without producing injury from excessive concentration. Theoretically this is feasible if no restriction is placed on the volume of nutrient solution in relation to the number of plants. In an experiment with tomatoes (Arnon & Hoagland, 1940) it was found that the aggregate absorption of K^+, NO_3^-, $H_2PO_4^-$, Ca^{++}, and Mg^{++} by one tomato plant for a period of 5 months was around 2600 m.equiv. It would have required about 80 l. of the nutrient solution used to furnish this supply of nutrients. For most purposes it is obviously more practicable to change or replenish a nutrient solution in the course of an experiment than to attempt to provide the total supply of nutrients at the start.

Nutrient solutions with variable and constant components. Studies in plant nutrition often call for the use of a series of nutrient solutions differing from each other in one or more components. With three-salt solutions various proportions of salts can be conveniently plotted on an equilateral triangle of three co-ordinates (Shive, 1915). This arrangement provides for a variation in concentration of one or more of the three component salts, keeping the others constant. It entails, however, a simultaneous change in the concentration of the cation when the anion is varied, and *vice versa*. Recently a modified arrangement of the triangular system has been proposed (Hamner, Lyon & Hamner, 1942) which provides for the variations in the concentrations of certain anions with the cations remaining the same.

Frequently the desire is to omit an ion entirely, or to vary its concentration over a considerable range. This can usually be accomplished by allowing for a variation in the concentration of one other ion. The slowly absorbed sulphate ion has served as the balancing ion to maintain the constancy of composition of nutrient solutions for purposes such as a comparison of the merits of ammonium versus nitrate as sources of nitrogen (Arnon, 1937), the effect of hydrogen-ion concentration from *p*H 3 to 9 on growth (Arnon & Johnson, 1942), and the effect of variable potassium and phosphate concentration on the composition of fruit and foliage (Arnon & Hoagland, 1943).

Preparation of the nutrient solution. The preceding discussion has already indicated that no one formula for a nutrient solution is inherently superior to all others. A comprehensive compendium of many representative formulae for nutrient solutions has been recently prepared by Livingston (1942). The preparation of a solution is conveniently divided into two parts: the supplying of six elements, sometimes designated as macronutrients, required by plants in relatively large amounts, potassium, calcium, magnesium, nitrogen, phosphorus, and sulphur, and six micronutrient elements required in minute quantity, iron, manganese, boron, copper, zinc and molybdenum. The macronutrients represent three cations and three anions and can thus be furnished most simply as solutions of three salts, for example, calcium nitrate, mono-potassium phosphate and magnesium sulphate. Very early, however, it was often found desirable to furnish the macronutrients from four salts (Knop, 1865), and recent studies on the proportions in which the elements are absorbed by the plant support the practice (Arnon & Hoagland, 1940). Each of the micronutrients is usually added in a concentration less than 1 p.p.m. as against 50–1000 p.p.m. for different macronutrients. For some plants, zinc, copper and molybdenum may not have to be deliberately added as they will often be adequately supplied as contaminants, but this is a point to be ascertained rather than taken for granted. Boron and manganese are generally added in a concentration of *c.* 0·5 p.p.m. each. Iron may be added at frequent and regular intervals in amounts equivalent to 0·5 p.p.m. A convenient way of supplying iron to sand culture is by an initial mixing of a small amount of powdered magnetite with the sand (Eaton, 1941); no further additions of soluble iron are then required.

A considerable number of nutrient solutions have total concentrations of salts equivalent to osmotic pressures ranging from 0·5 to 1 atm. This range seems to be well suited to the physiological tolerances of most plants, but some latitude is permissible in increasing the concentration of the solution. Plants

differ, however, in their salt tolerance and no generalizations seem warranted. For many species it is desirable to have at all times a slightly acid reaction of the solution, to avoid precipitation, but it is by no means necessary to maintain a constant *p*H. Many plants grow very well in solutions from *p*H 4·5 to 6·5. A solution can be 'physiologically buffered' within this range by the addition of a small part of its nitrogen in the form of NH_4^+ (Trelease & Trelease, 1935). The *p*H of several solutions prepared in this way demanded but little attention in extended experiments (Arnon & Hoagland, 1940). When necessary, adjustments may be made with sodium hydroxide and sulphuric or nitric acid.

Iron and other heavy metals can be made available at alkaline reactions if supplied as synthetic humate metals (Horner, Burk & Hoover, 1934). According to one report inorganic iron can also be kept in solution at alkaline reactions by the addition of small amounts of sodium metaphosphate (Edgerton, 1942). The importance of micronutrients, often present as impurities in nutrient salts or in water, or derived from containers, makes it imperative to exercise care as to the sources of chemicals and water for critical experiments, but ordinary tap water and technical salts, if relatively pure, may be employed in large-scale nutrient solution cultures for greenhouse crops in the investigation of general problems of nutrition.

III. THE USE OF ISOTOPES AND INDICATOR ELEMENTS

Whatever methods of artificial culture are adopted for the study of the inorganic nutrition of plants, problems arise as to the analytical determination of chemical elements undergoing absorption or translocation. Elements essential for growth are necessarily present in the plant *ab initio*. Increments of an element absorbed during an experiment must be computed on the basis of the initial status of the tissue. In various short-time experiments, in which only a small quantity of an ion is absorbed, unavoidable errors of plant variability or of laboratory technique may therefore render interpretation uncertain.

The use of a test ion not already found in the plant, or found to only a negligible extent, has advantages. In early work the problem presented itself in researches on certain large algal cells (Hoagland & Davis, 1923). To meet the need for a test element, bromide was successfully introduced. It is not toxic in the concentrations required, is readily mobile in the plant and has been profitably used as an indicator in numerous experiments with higher plants. Rubidium is another test element that has been used (Lundegårdh, 1936, 1940; Steward & Harrison, 1939). Its chemical determination in the presence of potassium is generally too difficult, but this and other elements can be determined spectroscopically and these methods have been developed and extensively employed by Lundegårdh (1936). Collander (1941) has found these spectroscopic methods valuable in studying the selective absorption of ions by different species of plants growing in the same nutrient solution.

An extremely sensitive and convenient method is to employ as tracers radioactive or stable isotopes. The dynamic character of protein synthesis and breakdown in the plant has been investigated by Vickery, Pucher, Schoenheimer & Rittenberg (1940) by using a stable isotope of nitrogen, following the extensive researches by animal biochemists. Some inorganic elements normally entering into plant metabolism, for example potassium and phosphorus, can be labelled by their radioactivity. The chemical properties are the same as those of the non-radioactive elements, but the radiation of the isotopes permits the detection of their movements with certainty, even when the amounts present are exceedingly minute. It is possible that the effects of radiation of sufficient intensity may injure or otherwise alter the protoplasm, but so far in most experiments on higher plants subjected to dilute culture solution evidence has not appeared that this action invalidated the results, although further observations are desirable. The isotope method sometimes permits of unique experiments. As an example, by ordinary chemical analysis the net movement of an element into root cells may be ascertained, while the concurrent outward movement of the same element, previously absorbed by the roots in radioactive form, can be segregated (Overstreet & Broyer, 1940). Radioactive isotopes, including those of essential elements, can be included in an ordinary culture solution and then over short time intervals the movement of the isotope can be observed. The translocation of inorganic nutrients from root to shoot, their accumulation in active tissues, including developing fruit and seed, and the relations of water movement to solute movement, have all been investigated by the new technique (Biddulph, 1941; Colwell, 1942; Stout & Hoagland, 1939; Arnon, Stout & Sipos, 1940; Gustafson and Darken, 1937; general review by Hevesy, 1940).

IV. ARTIFICIAL CULTURE MEDIA AND THE CONTROL OF THE AERIAL ENVIRONMENT

Since plant nutrition depends on the carbon-assimilating organs as well as on the utilization of inorganic nutrients absorbed by the roots, atmospheric control as well as control of the root medium is often essential to quantitative experiments. For complete control of the environment air-conditioned chambers illuminated with artificial light are requisite. The automatic regulation of a given air temperature or humidity is feasible, but the provision of illumination of suitable intensity and quality presents real difficulty. Imitation of intense sunlight is not now capable of achievement. Never-

theless important additions to knowledge gained by growing plants in artificial light have been recorded by numerous investigators in recent years (Steinberg, 1931; Lundegårdh, 1932; Trelease, 1925; Wilson, 1937; Brown, 1939; Arthur, Guthrie & Newell, 1930; Rohrbaugh, 1942; Greider & Downs, 1932; Crocker, 1925; Stoughton, 1930; Mitchell, 1936; Went, 1943; and others). The light from incandescent (Mazda) lamps concentrated on a small chamber is found to permit growth of wheat plants comparable with those grown under natural conditions (Davis & Hoagland, 1928), accomplishing a complete cycle with the production of normal seed. But many types of plants do not succeed so well with such illumination and often the growth is abnormal. For this reason investigators have frequently used other sources of light. Powerful arc lights, properly screened, give more or less satisfactory illumination for growth of some kinds of plants, at least in vegetative stages. Recently the new fluorescent lamps have attracted much attention and various combinations of colours can be made to vary the spectral distribution of the light energy. The problem of intensity is still an obstacle. Finally, carbon dioxide concentration in the atmosphere should be mentioned as a factor. This is usually assumed to be controlled at atmospheric concentration, but experiments are also sometimes carried out with enhanced concentrations (Bolas & Melville, 1933).

In some investigations artificial root media are used with partial control of the aerial environment, particularly humidity or temperature, or control of illumination by subjecting plants to varying photo periods. Thomas, Hendricks, Collier & Hill (1943) have studied the growth of crops under natural sunlight in glass houses, with the control of root media by sand cultures. Ingenious automatic measurements of photosynthesis or respiration are made and temperature and light intensity, although varying with natural conditions, are also recorded. By such methods one could assess climatic influences in their interrelations with supplies of nutrients in the soil.

V. MICRONUTRIENT ELEMENTS

The older teachings of plant physiology were erroneous in limiting the list of chemical elements essential for growth of higher plants to the following ten: carbon, hydrogen, oxygen, nitrogen, sulphur, calcium, magnesium, potassium, phosphorus and iron. Not less essential than these elements are minute quantities of boron, manganese, copper, zinc, and, according to recent evidence, probably molybdenum (Arnon & Stout, 1939; Brenchley, 1943). The significance of these elements for plant growth is greatly enhanced by the observations of the past ten years that deficiencies of boron, copper, manganese, and zinc account for numerous previously obscure plant diseases occurring in the field. Frequently the deficiencies may be overcome by application of the deficient element to the soil or directly to the plant, and the use of these curative measures in appropriate cases is of great agricultural importance. Long before the essential nature of these elements was clearly recognized plant physiologists were able to grow numerous species of plants in water culture with the deliberate addition to the culture solution of only seven elements: namely, nitrogen, sulphur, phosphorus, calcium, magnesium, potassium and iron, for the reason that the impurities in salts, in water, or from the vessels often supplied the necessary minute amounts of other essential elements, of the order of a few thousandths to a few tenths mg. l. Clearly, to prove that these elements are physiologically essential, the nutrient medium must be purified to such a degree that one element at a time can be eliminated, or so nearly eliminated, that the quantity still remaining is inadequate for physiological functions of the plant. The fact that a plant grows normally in a solution made of purified salts used to supply only the twelve elements now regarded as essential does not in itself warrant a conclusion that other elements are unessential. One can only assert that if other elements are essential, the quantities required by the plant under study are not greater than those remaining in the solution, or gaining access to it after taking the measures currently available for eliminating impurities. It is entirely possible, and in fact probable, that further improvements in water-culture technique may disclose additional essential elements. The fortunate selection of a species of plant that may be found to have a relatively high requirement for specific elements in microquantities will enable more rapid progress to be made in the discovery of essential elements. Storage of the elements in the seed, however, limits the degree of deficiencies to which the plant is subjected.

Because of the rigorous precautions that must be taken in purification of media and because of differential quantitative requirements by different plants, evidence for the essential nature of a particular element is not always easily secured and occasionally investigators express doubt that all the micronutrient elements mentioned above are indispensable for all species of higher plants. It is not, of course, feasible to settle these doubts by direct experimentation with every kind of plant, but the complete failure of plants of many diversified species tested to develop a normal cycle of growth, or often to make appreciable growth at all, in the absence of one of the micronutrient elements, strongly suggests that these elements perform fundamental roles in metabolism of all higher plants. In some cases failure to demonstrate the need of a micronutrient element is to be ascribed to inadequate technique. For example, ordinary distilled water sometimes contains enough copper as an impurity to insure the necessary supply of this element for plant growth. Convenient methods for purifying nutrient salts of certain metals have been recently described by Steinberg (1938)

and later by Stout & Arnon (1939), who have shown that it is possible to reduce the aggregate heavy metal impurity of the culture solution to less than one part per billion. Other techniques, such as recrystallization of salts or preparation of salts from original elemental constituents, were described by Sommer & Lipman (1926) in some of the pioneering studies. These investigators obtained clear evidence of the essential nature of several micronutrients, following the report on experiments with purified salt solutions by Mazé (1914). Early studies on boron essentiality were made by Brenchley & Warington (1927). Many later investigators have adopted the water-culture technique for micronutrient studies. Others have selected the sand-culture technique which involves the purification of sand. Thus Hageman, McHargue, Sherman & Hodge (1924), studied manganese deficiency in oats, while McHargue (1922) had early obtained evidence for the essential nature of manganese. In general, however, the preparation of sand free of impurities is fraught with greater difficulty than the preparation of a purified nutrient solution, and it is not likely that our present status of knowledge could have been attained without the water-culture technique. The animal physiologist meets greater difficulty in investigation of micronutrient elements because the preparation of a diet free of inorganic impurities is far less readily accomplished.

To determine which chemical elements are essential for plant growth in minute quantity is only an initial step in the whole investigation, and the usefulness of artificial culture methods does not end here. The most interesting problems are those of the functions of the elements in metabolism. Of these functions relatively little is known, but artificial culture methods now make more readily feasible the growing of plants with graduated supplies of a micronutrient element. The use of water or sand-culture methods for this purpose are found in studies by Skoog (1940), Reed (1941), and Reed & Dufrenoy (1942) on zinc, and by Shive and his collaborators on boron (review by Shive, 1941). Experiments with plants under artificial conditions can also furnish valuable information on micronutrient elements in their relation to agriculture. The symptoms in the plant of a controlled deficiency are observable and serve as a guide when field tests leave uncertainty whether the beneficial effect of supplying a given element to diseased plants is the direct response of removing a deficiency or an indirect response such as might arise from neutralization of a toxic substance. Moreover, micronutrient elements become toxic when very low concentrations in the medium are exceeded. In certain arid regions boron toxicity is of practical significance. By artificial culture methods relative tolerances of different species of plants are conveniently studied. Extensive sand-culture experiments have been carried out on boron toxicity (Eaton, 1935).

VI. ROOT ABSORPTION AND UPWARD MOVEMENT OF INORGANIC NUTRIENTS

A fundamental aspect of plant nutrition is the absorption by root cells of salts or ions and their movement to the upper parts of the plant. These phenomena can be studied on plants growing in soils and for many purposes this procedure is essential, but an artificial culture technique is advantageous, or indeed sometimes indispensable. By immersing the root system in an aqueous inorganic solution the difficulties of appraising a system of liquid and solid phases is avoided. What the plant has removed can be determined by analysis of the culture solution, although it is often necessary to analyse the plant tissues or the sap expressed from them. The relation of water absorption to salt absorption is of significance and the nutrient-solution technique is well adapted to making pertinent comparisons. Experiments demonstrate that there is no necessary correspondence between the quantities of solutes and of water absorbed. Some ions may be absorbed relatively more rapidly than water, and some less rapidly (Hoagland, 1940; Broyer & Hoagland, 1943). By similar methods observations are readily made on the differential absorption of ions of a single salt. For example, the result has frequently been obtained that an actively absorbing plant removes from a solution of potassium sulphate more potassium ions than sulphate ions. The solution then becomes more acid, even to the point of injury to the root system. The elaboration of experiments of this type provides a basis for consideration of many aspects of salt absorption from the point of view of ionic exchanges between the root and the culture medium.

These and other effects of the plant on the nutrient medium, the effects of one ion on the absorption by the plant of another ion simultaneously present in the nutrient solution, and concentration relations, may all be examined with special convenience by the use of appropriate artificial methods, especially the liquid-culture method, but of greater significance are the opportunities afforded by these methods for research on the metabolism of plants as influencing the intake of ions by the roots. Much work has been done in recent years with young wheat and barley plants (Hoagland & Broyer, 1940; Prevot & Steward, 1936; Lundegårdh, 1940). The plants are grown in a nutrient solution for a few weeks and then subjected to experimentation often limited to a few hours. By this means the investigator is able to compare the responses of uniform sets of plants to different solutions and to other different environmental conditions governing the metabolic activities of roots or shoots.

While the plant with its roots immersed in a solution forms a much simpler system than a plant growing in a soil, a further simplification can be introduced by excising the root systems and investigating the absorption of ions or the metabolism of

the isolated roots. The excised roots have at their disposal organic metabolites translocated from the green parts of the plant during the initial growth period and these suffice to maintain active root metabolism over short intervals. The nutrient status of the plant at the beginning of the experiment can be modified by growing plants under different nutrient conditions in the culture solution, or by varying the light and temperature conditions around the shoots (Hoagland & Broyer, 1936). Another culture technique is provided by growing seedlings in a nutrient solution for only a few days in a chamber with temperature control, with or without illumination. Subsequently single unbranched roots are excised and subdivided. Thus small amounts of living root tissues are available for study in a microrespirometer. Parallel experiments can be made on the absorption of inorganic solutes, including radioactive isotopes (Machlis, unpublished).

The methods outlined above are suitable for work on the physiology of the plant itself, but for understanding soil-plant interrelations there is need also for research on colloidal systems. In a soil, absorbing roots develop intimate contacts with colloidal particles, and it becomes desirable to devise experiments in which colloids are present, without incurring all the difficulties presented by the entire complex medium of the soil. One technique for this purpose is based on the use of solutions, or of pure water, in which purified soil colloids are suspended. From such experiments, Jenny & Overstreet (1939) suggest that ions are not necessarily always directly absorbed by roots from a solution in accordance with a soil solution mechanism, but that an additional mechanism exists. This envisages essentially a direct transfer of ions from one colloidal system to another, within the oscillation volumes of the adsorbed ions. The soil colloids especially concerned hold positively charged ions, which can exchange for the positively charged ions held on root surfaces through zones of contact. Thus hydrogen ions metabolically generated by the plant can be exchanged for potassium, calcium, magnesium, or sodium ions held by the soil colloid, and the plant gains in these ions. On the other hand, if the soil colloid is sufficiently near saturation by hydrogen or sodium ions, the so-called 'contact depletion' effect may occur, by which calcium, magnesium, or potassium ions previously absorbed are lost by the roots to the soil colloid by ionic exchange. These initial steps of ion exchanges do not of course explain the subsequent events in which ions move inward under the influence of metabolic processes. The relative importance in the soil of absorption of ions by plants through a contact mechanism as compared with absorption from the soil solution is not yet clear. Nitrate and certain other ions appear to be almost exclusively present in the soil solution. The adoption of an artificial medium and the application of radioactive isotopes supply tools of research for direct experimentation on the contact theory, which, while having as a special objective the nature of soil-plant interrelations, cannot be readily examined by direct analysis of a soil-plant system. From the point of view of soil and plant interrelations, ion exchange phenomena are of such interest that various investigators developed artificial culture media from which plants may derive adsorbed cations and anions. Recently developed cation and anion exchange compounds make feasible this new approach (Converse, Gammon & Sayre, 1943; Schlenker, 1940; Graham & Albrecht, 1943).

It is not feasible here to review the literature of salt absorption and accumulation which has been done elsewhere. This includes researches on storage tissues and algal cells, which have been so prominent in work on this aspect of plant nutrition. There may be some value, however, in indicating the general nature of the processes of salt accumulation by plant roots, since much of the information is derived from experimentation based on artificial culture technique. If the roots of the plant are placed in a culture solution, the removal by the plant of salts (or ions) does not normally proceed by a simple diffusion mechanism, in which ions enter the roots according to positive concentration or activity gradients. If the external solution is dilute, ions may move inward from a low external to a higher internal concentration, as shown by studies on root sap. This movement against a gradient is also evident from root exudates, which show a polarized migration of solutes from root cells into the conducting system of the plant. These types of solute movements demand a source of metabolic energy, intimately dependent on aerobic respiration and other accompanying metabolic processes (Hoagland & Broyer, 1940; Prevot & Steward, 1936; Lundegårdh, 1940). A controlled artificial culture technique likewise offers a means for the study of the respiratory processes in relation to salt accumulation, through observation on respiratory poisons or inhibitors. By direct biochemical studies on roots grown in culture solutions preliminary evidence has been obtained on the metabolism of carbohydrate or nitrogen compounds. This problem is one of wide interest, and researches on other tissues, especially storage ones, have shown the importance of aerobic metabolism in relation to salt absorption (reviews by Steward, 1935–7).

VII. AGRONOMIC AND HORTICULTURAL PROBLEMS

Many studies by artificial culture methods have been conducted to gain supplementary information on specific problems of practical agriculture associated with fertiliser practices. Here we can only refer to some types of questions which lend themselves to investigation by artificial culture technique. The recognition in specific crops of deficiency symptoms in nutrient elements required in larger as well as in minute amounts is the goal of some experiments.

Comparison may be made of plants grown under the artificial conditions with plants grown in soil. This procedure serves useful purposes, although the interreactions of many factors often precluded the certain diagnosis of a deficiency occurring in the field on the basis of symptomatic response alone. A handbook of pathological symptoms caused by nutrient deficiencies in numerous species of agricultural plants grown in soil or in artificial culture is now available (American Society of Agronomy and National Fertiliser Association).

Even when a deficiency of a nutrient element is not so acute as to result in marked pathological disturbances, a nutrient factor may limit crop yields, and no question recurs more frequently than that of attempting to learn, apart from laborious field experiments, what fertilisers applied to a soil will give profitable increases in crop production or improvement in quality. Immense efforts have been made to obtain guidance by chemical analyses of the soil. Often these efforts are not attended with great success for reasons that cannot be discussed here. Another approach is to appeal to the plant and analyse its tissues instead of the soil. When this is done the underlying assumption is that the plant is the living integrator of the entire system of soil, plant and atmosphere. The method is far from new, yet in recent years a recrudescence of interest in the plant diagnostic procedure is evident. Again the artificial culture technique finds an application, in that fluctuations in plant composition that can be induced by wide variations in nutrient supply are more readily investigated and interpreted than by observations on plants growing in a soil.

The extraordinary importance of soil treatments with nitrogen fertilisers is familiar to all agriculturists. The effects of nitrogen supply, however, cannot be properly evaluated without reference to the relations of nitrogen metabolism to carbon assimilation and metabolism, particularly when the harvestable parts of the crop consist of fruit or seed. Reactions of certain carbon compounds with ammonia nitrogen are rapid in growing plants. While older ideas of carbon-nitrogen relations affecting fruitfulness are undergoing modification in the light of new knowledge of plant hormones, the importance of the metabolic reactions of carbon compounds and inorganic nitrogen compounds is still maintained. To understand phenomena of this kind there is great value in artificial culture methods, which permit control of the supply of nitrogen at different stages of plant development, and under different atmospheric environments. The interreactions of nutrient elements other than nitrogen in the metabolic system can of course be studied by similar methods. Gregory and his associates (1937) describe experiments by sand-culture methods on nutrient deficiencies affecting barley. Sand culture and water culture have been employed on nutritional studies on fruit trees by various investigators (Wallace, 1933; Colby, 1933). Irrespective of practical agriculture, the biologist finds a field for artificial media in investigating plant metabolism. Interest in this field is augmented by comparisons of the metabolism of plant and animal, as in discussions by Chibnall (1939) and by Vickery (1937) and his collaborators on the nitrogen and organic acid metabolism of plants.

VIII. THE EFFECTS OF 'ALKALI' ON PLANTS

Investigators of agricultural subjects are familiar with problems of arid or semi-arid regions, associated with accumulation of soluble salts in the soil, especially sodium salts. There are many aspects of these problems. One of them is that of the tolerance of various species of plants to high concentrations of salt or to conditions resulting from the action of sodium salts on soil colloids. The appraisal of effects of alkali on plants growing in soil is subject to vast complications and certain points can be most securely elucidated by artificial culture experiments (Reed & Haas, 1923; Lipman, Davis & West, 1926; Heller, Hageman & Hartman, 1940). Both water culture and sand culture provide ways of learning more of the response of plants to controlled root environments with respect to salt concentrations. The U.S. Federal Government has established at Riverside, California, a laboratory for the study of the alkali problem, and included in the programme of research are experiments with artificial cultures (Magistad, Ayers, Wadleigh & Gauch, 1943; Hayward & Long, 1942; Eaton, 1942; Gauch & Eaton, 1942).

IX. COMMERCIAL USE OF ARTIFICIAL CULTURE METHODS

During most of the long history of artificial culture of higher plants the central objective has been either to advance our knowledge of plant physiology, or to conduct experiments for supplementing other information on the growth of plants in soils. In recent years another objective has been much discussed, and has aroused an astonishing interest on the part of many not engaged in scientific pursuits. The objective referred to is to employ artificial culture methods to grow crops commercially. The general principles of the production of plants by artificial culture methods for commercial purposes are the same as those for scientific experiments. The scale, however, is usually different, and, above all, the cost of production enters as a factor. Save in unusual circumstances, the method is commercially feasible mainly for greenhouse crops. The value of space and the selling price of the crop must be high to warrant the soil medium being displaced by artificial culture. The gain in control of the nutrition of the plant and the avoidance of infections or pests arising from the

use of soil in a greenhouse are often cited as reasons for adopting an artificial culture method as a new greenhouse technique. The use of such media offers no short cut to successful growing of crops and all the general knowledge and experience required for the production of a specific greenhouse crop will still be essential and, in addition, an understanding of new techniques in the management of the nutrient medium.

Sometimes there has been a popular impression that enormously larger yields can be obtained per unit of surface from certain artificial media than from soil. Actually, climatic and genetic factors, and, often, plant diseases and pests, limit crop production whether a favourable soil or a favourable artificial medium is employed. Furthermore, requirements for illumination govern the density of spacing even when an artificial medium could supply more nutrients than the soil. Obviously, if a poor or mediocre soil is compared with a highly suitable artificial nutrient medium, the crop-producing power of the latter may be expected to surpass greatly that of the soil. A proper comparison of two types of media demands that each be maintained at or near its optimum, and that crops be compared under otherwise identical conditions. There is but little value in contrasting crops produced by artificial culture in a greenhouse with yields under average field conditions.

An experiment was conducted in California in which different lots of tomato plants were grown in the same greenhouse, some in sand, some in water culture, and some in a soil heavily fertilized and treated with organic matter. The yields for all three methods were exceptionally high, and of the same magnitude. But it appeared that the yields from the sand and from the water culture, provided the latter received forced aeration, were considerably larger than from the soil (Arnon & Hoagland, 1940). The explanation has been suggested that in the sand culture and in the forcibly aerated water culture the aeration of the roots approached the optimal more closely than was possible in a thickly planted bed of soil. Tomato plants have a high aeration requirement. On the other hand, Templeman & Watson (1938) reported an experiment in which the yields of tomato plants in soil and artificial media were compared. Their preliminary results did not show a superior yield from the artificial media, but rather the reverse.

The three general methods of artificial culture tried in commercial practice are sand, gravel and water culture. One special adaptation of the water-culture method devised by Gericke (1940) has been named 'hydroponics'. This technique includes the use of large tanks of nutrient solution over which are placed porous beds. Some root development takes place in the beds as well as in the solution. While the porous-bed method facilitates aeration of roots, in the experiment cited above additional forced aeration of the nutrient solution was beneficial. The porous-bed technique, as well as climatic influences, are stressed in the report by Gericke. It is too early to evaluate the place of artificial culture methods in crop production even in greenhouse practice. In America the gravel-culture method is probably receiving most attention by commercial greenhouse operators. Unfortunately no adequate statements seem to be available reporting failures and difficulties in practice, as well as successes, for various systems of artificial culture together with comparisons of costs and profits (General discussion, Hoagland & Arnon, 1938).

X. ANIMAL NUTRITION IN RELATION TO PLANT NUTRITION

Since the green plant is the primary synthetic unit from which the animal derives its sustenance, students of animal nutrition have an immediate concern with the composition of the plant as influenced by environmental factors. A systematic effort to build up a working knowledge of the relations of soil, plant nutrition and animal nutrition was inaugurated several years ago in the U.S.A. in a new Federal laboratory at Cornell University. Other institutes in Australia, Great Britain, and elsewhere have recognized by some of their researches the significance of the soil-plant system to the nutrition of the animal. Obviously, many of the objectives in this field of study must be gained by the chemical analysis of plants grown in soil, or by animal feeding trials with soil-grown plants. These are methods that have succeeded in showing that plants grown under certain soil conditions are deficient in quality from the point of view of animal nutrition. Deficiencies of iodine, cobalt, iron, calcium, phosphate, or other inorganic elements have received attention. These are usually cases in which extremes of deficiency exist and so their recognition rests on their exceptional nature. The enormous extent of present research on animal nutrition and the development of understanding of vitamins, essential amino-acids and other factors of animal nutrition now justify much more systematic and intensive probing into the nutrition of plants in its bearing on the problems of animal nutrition. The relative effects of soil and climate on the elaboration of vitamins by the plant, represents one type of inquiry. The difficulty of answering such questions will be apparent and controlled artificial culture methods have an important place in a research programme. A recent example is found in a test of vitamin C content of tomatoes of the same strain grown in several climatic locations both in soil and in culture solutions (Hamner, Lyon & Hamner, 1942).

XI. SUMMARY

This review presents a general survey of artificial culture methods employed in the investigation of plant nutrition. Several types are described, including liquid-culture

methods and those which depend on a solid inert medium. The advantages of artificial culture procedures for growing plants are pointed out, as one means of studying soil-plant interrelations, as well as various questions in plant physiology and plant biochemistry. These methods have been indispensable in the study of the chemical elements essential to the growth of higher plants, especially of those elements needed in minute quantities. The techniques of artificial culture are also of great value for the investigation of the absorption of ions by roots. These techniques serve likewise for inquiries into the interreactions of climatic conditions and mineral nutrients. Among other topics considered are the application of artificial culture methods to researches on the functions in plant metabolism of inorganic nutrients, the role of colloids in absorption of ions, horticultural and agronomical problems, and commercial production of crops.

XII. REFERENCES

AMER. SOC. AGRON. AND NAT. FERT. ASS. (1941): *Hunger Signs in Crops.*

ARNON, D. I. (1937): *Soil Sci.* **44**, 91. — (1938): *Amer. J. Bot.* **25**, 322. — (1943): *Ann. Rev. Biochem.* **12**, 493.

ARNON, D. I. & HOAGLAND, D. R. (1940): *Soil Sci.* **50**, 463. — (1943): *Bot. Gaz.* **104**, 576.

ARNON, D. I. & JOHNSON, C. M. (1942): *Plant Physiol.* **17**, 525.

ARNON, D. L. & STOUT, P. R. (1939): *Plant Physiol.* **14**, 599.

ARNON, D. I., STOUT, P. R. & SIPOS, F. (1940): *Amer. J. Bot.* **27**, 791.

ARTHUR, J. M., GUTHRIE, J. D. & NEWELL, J. M. (1930): *Amer. J. Bot.* **17**, 416.

BIDDULPH, O. (1941): *Amer. J. Bot.* **28**, 348.

BIEKART, H. M. & CONNORS, C. H. (1935): *Bull. N.J. Agric. Exp. Sta.* no. 588, p. 3.

BLAKE, M. A., NIGHTINGALE, G. T. & DAVIDSON, O. W. (1937): *Bull. N.J. Agric. Exp. Sta.* no. 626, p. 1.

BOLAS, B. D. & MELVILLE, R. (1933): *Ann. Bot., Lond.*, **47**, 673.

BOUSSINGAULT, J. (1886): *Recherches sur la Végétation*, 3rd ed. Paris.

BRENCHLEY, W. E. (1936): *Bot. Rev.* **2**, 173. — (1943): *Biol. Rev.* **18**, 159.

BRENCHLEY, W. E. & WARINGTON, K. (1927): *Ann. Bot., Lond.*, **41**, 167.

BROWN, E. M. (1939): *Plant Physiol.* **14**, 517.

BROYER, T. C. & HOAGLAND, D. R. (1943): *Amer. J. Bot.* **30**, 261.

CHAPMAN, H. D. & BROWN, S. M. (1941): *Hilgardia*, **14**, 161. — (1941): *Hilgardia*, **14**, 185.

CHAPMAN, H. D. & LIEBIG, Jr., G. F. (1938): *J. Agric. Res.* **56**, 73. — (1940): *Hilgardia*, **13**, 141.

CHIBNALL, A. C. (1939): *Protein Metabolism in the Plant.* Yale Univ. Press.

COLBY, H. L. (1933): *Plant Physiol.* **8**, 1.

COLLANDER, R. (1941): *Plant Physiol.* **16**, 691.

COLWELL, R. N. (1942): *Amer. J. Bot.* **29**, 798.

CONVERSE, C. D., GAMMON, N. & SAYRE, J. D. (1943): *Plant Physiol.* **18**, 114.

CROCKER, W. (1925): *Plant Res. Boyce Thompson Inst.* **1**, 9.

DAVIS, A. R. & HOAGLAND, D. R. (1928): *Plant Physiol.* **3**, 277.

DUHAMEL DU MONCEAU, H. L. (1758): *La Physique des Arbres*, **2**. Paris.

DUNLAP, A. A. (1936): *Conn. Agric. Exp. Sta. Bull.* no. 380, p. 136.

EATON, F. M. (1931): *Soil Sci.* **31**, 235. — (1935): *Tech. Bull. U.S. Dep. Agric.* no. 448. — (1936): *J. Agric. Res.* **53**, 433. — (1941): *Plant Physiol.* **16**, 385. — (1942): *J. Agric. Res.* **64**, 357.

EDGERTON, L. J. (1942): *Proc. Amer. Soc. Hort. Sci.* **41**, 237.

FURNSTAL, A. H. & JOHNSON, B. (1936): *Plant Physiol.* **11**, 189.

GAUCH, H. G. & EATON, F. M. (1942): *Plant Physiol.* **17**, 347.

GERICKE, W. F. (1940): *The Complete Guide to Soilless Gardening.* Prentice Hall, N.Y.

GILBERT, S. G. & SHIVE, J. W. (1942): *Soil Sci.* **53**, 143.

GRAHAM, E. R. & ALBRECHT, W. A. (1943): *Amer. J. Bot.* **30**, 195.

GREGORY, F. G. (1937): *Ann. Rev. Biochem.* **6**, 557.

GREGORY, F. G. & CROWTHER, F. (1928): *Ann. Bot., Lond.*, **42**, 757.

GREIDER, C. E. & DOWNES, A. C. (1932): *Trans. Illum. Eng. Soc., N.Y.*, **27**, 637.

GUSTAFSON, F. G. & DARKEN, M. (1937): *Amer. J. Bot.* **24**, 615.

HAGEMAN, R. H., MCHARGUE, J. S., SHERMAN, D. G. & HODGE, E. S. (1942): *J. Amer. Soc. Agron.* **34**, 731.

HAMNER, K. C., LYON, C. B. & HAMNER, C. L. (1942): *Bot. Gaz.* **103**, 586.

HAYWARD, H. E. & LONG, E. M. (1942): *Proc. Amer. Soc. Hort. Sci.* **41**, 149.

HELLER, V. G., HAGEMAN, R. H. & HARTMAN, E. L. (1940): *Plant Physiol.* **15**, 727.

HEVESY, G. (1940): *Ann. Rev. Biochem.* **9**, 641.

HOAGLAND, D. R. (1923): *Soil Sci.* **16**, 225. — (1940): *Cold Spr. Harb. Symp.* **8**, 181.

HOAGLAND, D. R. & ARNON, D. I. (1938): *Circ. Calif. Agric. Exp. Sta.* no. 347, p. 1.

HOAGLAND, D. R. & BROYER, T. C. (1936): *Plant Physiol.* **11**, 471. — (1940): *Amer. J. Bot.* **27**, 173.

HOAGLAND, D. R. & DAVIS, A. R. (1923): *J. Gen. Physiol.* **6**, 47.

HOAGLAND, D. R. & SNYDER, W. C. (1933): *Proc. Amer. Soc. Hort. Sci.* **30**, 288.

HORNER, C. K., BURK, D. & HOOVER, S. R. (1934): *Plant Physiol.* **9**, 663.

JENNY, H. & OVERSTREET, R. (1939): *Soil Sci.* **47**, 257.

JOHNSTON, E. S. (1927): *Plant Physiol.* **2**, 213.

JOHNSTON, E. S. & HOAGLAND, D. R. (1929): *Soil Sci.* **27**, 89.

KNOP, W. (1862): *Landw. VersSta.* **4**, 173. — (1865): *Landw. VersSta.* **7**, 93.

LAURIE, A. & KIPLINGER, D. C. (1940): *Bull. Ohio Agric. Exp. Sta.* no. 616, p. 1.

LIEBIG, JUSTUS VON (1840): *Die organische Chemie in ihrer Anwendung auf Agrikultur und Physiologie.* Braunschweig.

LIPMAN, C. B., DAVIS, A. R. & WEST, E. S. (1926): *Soil Sci.* **22**, 303.

LIPMAN, C. B. & MACKINNEY, G. (1931): *Plant Physiol.* **6**, 593.

LIVINGSTON, B. E. (1942): *Publ. Amer. Ass. Adv. Sci.* **16**, 83.
LUNDEGÅRDH, H. (1932): *Die Nährstoffaufnahme der Planze.* — (1936): *Lantbrukshögskolan Annaler*, **3**, 49. — (1940): *Ann. Agric. Coll. Sweden*, **8**, 233.
MCHARGUE, J. S. (1922): *J. Amer. Chem. Soc.* **44**, 1592.
MAGISTAD, O. C., AYERS, A. D., WADLEIGH, C. H. & GAUCH, H. G. (1943): *Plant Physiol.* **18**, 151.
MAZÉ, P. (1914): *Ann. Inst. Pasteur*, **28**, 21.
MITCHELL, J. W. (1936): *Plant Physiol.* **11**, 833.
NIKLAS, H. & SCHROPP, W. (1931): *Fortschr. Landw.* **12**, 1.
OVERSTREET, R. & BROYER, T. C. (1940): *Proc. Nat. Acad. Sci., Wash.*, **26**, 16.
PARKER, F. W. & PIERRE, W. H. (1928): *Soil Sci.* **25**, 337.
PREVOT, P. & STEWARD, F. C. (1936): *Plant Physiol.* **11**, 509.
RAINES, N. A. (1937): *Amer. J. Bot.* **24**, 185.
REED, H. S. (1941): *Amer. J. Bot.* **28**, 10.
REED, H. S. & DUFRENOY, J. (1942): *Amer. J. Bot.* **29**, 544.
REED, H. S. & HAAS, A. R. C. (1923): *Tech. Pap. Agric. Exp. Sta.* no. 4, p. 1.
ROHRBAUGH, L. M. (1942): *Bot. Gaz.* **104**, 135.
SACHS, JULIUS VON (1887): *Lectures on the Physiology of Plants.* Oxford: Clarendon Press.
SALM-HORSTMAR, F. (1856): Diss. Braunschweig.
SCHLENKER, F. S. (1940): *Amer. J. Bot.* **27**, 525.
SHIVE, J. W. (1915): *Amer. J. Bot.* **2**, 157. — (1915): *Physiol. Res.* **1**, 327. — (1941): *Plant Physiol.* **16**, 435.
SHIVE, J. W. & ROBBINS, W. R. (1938): *Bull. N.J. Agric. Exp. Sta.* no. 636, p. 3.
SHIVE, J. W. & STAHL, A. L. (1927): *Bot. Gaz.* **84**, 317.
SKOOG, F. (1940): *Amer. J. Bot.* **27**, 939.
SOMMER, A. L. (1931): *Plant Physiol.* **6**, 339.
SOMMER, A. L. & LIPMAN, C. B. (1926): *Plant Physiol.* **1**, 231.
STEINBERG, R. A. (1931): *J. Agric. Res.* **43**, 1071. — (1938): *J. Agric. Res.* **57**, 461.
STEWARD, F. C. (1935): *Ann. Rev. Biochem.* **4**, 519. — (1937): *Trans. Faraday Soc.* **33**, 1006.
STEWARD, F. C. & HARRISON, J. A. (1939): *Ann. Bot., Lond.*, **3**, 427.
STOUGHTON, R. H. (1930): *Ann. Appl. Biol.* **17**, 90.
STOUT, P. R. & ARNON, D. I. (1939): *Amer. J. Bot.* **26**, 144.
STOUT, P. R. & HOAGLAND, D. R. (1939): *Amer. J. Bot.* **26**, 320.
TEMPLEMAN, W. G. & WATSON, S. J. (1938): *J. Minist. Agric.* **45**, 771.
THOMAS, M. D., HENDRICKS, R. H., COLLIER, T. R. & HILL, G. R. (1943): *Plant Physiol.* **18**, 345.
TOTTINGHAM, W. E. (1914): *Physiol. Res.* **1**, 133.
TRELEASE, S. F. (1925): *Bull. Torrey Bot. Cl.* **52**, 389.
TRELEASE, S. F. & LIVINGSTON, B. E. (1922): *Sci.* **55**, 483.
TRELEASE, S. F. & THOMSON, J. R. (1935): *Sci.* **81**, 204.
TRELEASE, S. F. & TRELEASE, H. M. (1935): *Amer. J. Bot.* **22**, 520.
VICKERY, H. B., PUCHER, G. W., SCHOENHEIMER, R. & RITTENBERG, D. (1940): *J. Biol. Chem.* **135**, 531.
VICKERY, H. B., PUCHER, G. W., WAKEMAN, A. J. & LEAVENWORTH, C. S. (1937): *Bull. Conn. Agric. Exp. Sta.* no. 399, p. 757.
WALLACE, T. (1933): *Imp. Bur. Fruit Prod. Tech. Comm.* no. 4, p. 1.
WEAN, R. E. (1935): *Sci.* **82**, 336.
WENGER, R. (1941): *Plant Physiol.* **16**, 621.
WENT, F. W. (1943): *Amer. J. Bot.* **30**, 157.
WILSON, A. R. (1937): *Ann. Appl. Biol.* **24**, 911.
WITHROW, R. P. & BIEBEL, J. P. (1936): *J. Agric. Res.* **53**, 693. — (1938): *Bull. Purdue Agric. Exp. Sta.* no. 232, p. 1.
WITHROW, R. B., BIEBEL, J. P. & EASTWOOD, T. M. (1943): *Circ. Indiana Agric. Exp. Sta.* no. 277, p. 1.
WOODWARD, J. (1699): *Philos. Trans.* **21**, 382.
ZINZADZE, CH. (1935): *Sci.* **81**, 540.
ZURBICKI, Z. (1933): *Plant Physiol.* **8**, 553.

EXPERIMENTAL ANALYSIS OF THE ASSOCIATION BETWEEN INVERTEBRATES AND UNICELLULAR ALGAE

By C. M. YONGE, University of Bristol

(*Received* 11 *December* 1943)

Unicellular algae, broadly divisible into green 'zoochlorellae' and brown 'zooxanthellae', occur in the tissues of great numbers of aquatic invertebrates, notably in the Protozoa, Porifera, Coelenterata, Turbellaria, Gastropoda and Lamellibranchia, and also, in more isolated cases, in Annelida, Rotifera and ascidians; in addition, possibly also in Polyzoa and Echinodermata. Buchner has summarized and discussed the literature, much of it descriptive, up to 1930; here attention is confined to more recent work, only older work which is experimental being directly quoted. Particular attention is paid to conditions in marine invertebrates. The origin, nature and effects of this 'symbiosis' between animals and plants vary widely in different cases as briefly outlined on two previous occasions (Yonge, 1934, 1935), and the present article is an expansion of the views there expressed and also contains a short discussion on the possible role of this 'imprisoned phytoplankton' in tropical seas where zooxanthellae are conspicuously abundant.

I. ORIGIN OF THE ASSOCIATION

There is a striking correspondence between the presence of associated algae and the occurrence of intracellular digestion (Yonge, 1934), the exceptions being the isolated cases of algae occurring in ascidians, Polyzoa, Echinodermata and Annelida. The presence of algae in certain compound ascidians of tropical waters has been described by Smith (1935) who found them present in the cloacal cavities except in *Didemnum viride* where they are embedded in the test. The animals live in waters infested with Protozoa and larvae containing zooxanthellae, and Smith concludes that, when animals are digested, some of the algae may pass through the gut and accumulate in the cloacal cavities where the excrement from the ascidians would provide inorganic food salts. The presence of algae in the common test indicates early entrance before this is properly formed, but in no case are the algae actually in the living tissues of the animal. In the Polyzoa, algae have been reported in *Zoobothryon* and *Bicillaria*, but in neither case has Buchner (1930) been able to confirm this. In the same way the isolated cases of their presence in Echinodermata require confirmation (see Buchner); in many cases they may be pigment cells. Berkeley (1930*a*) has described a clear case of association in Annelida. In species of three genera of the Chaetopteridae the intestinal wall contains numerous green unicellular organisms, 1–3 μ in diameter, which, however, possess 'neither cellulose wall, cell wall, chromatophore nor pyrenoid and appear to contain no algal pigment'. He regards them as flagellates probably belonging to the Chrysocapsinae. He is unable to decide whether they are symbionts or parasites, but is certain that they must infect the animal early in development, Enders (1909) having already shown that they first appear in the larvae. The work of Wilson (1932), who found indication of intracellular digestion during the metamorphosis of the annelid *Owenia*, provides a possible explanation of the origin of the association. Berkeley (1930*b*) has also described association between the ctenophore *Beroe abyssicola* and very similar flagellates, red in colour, which occur in the subepithelial layers of the stomodaeum. This ctenophore lives in deep water, and Berkeley stresses the possible connexion, both here and in the Chaetopteridae, between the presence of algae and the occurrence of bioluminescence.

Undoubted cases of parasitic infection by species of the fresh-water alga *Chlorella* have been described by Goetsch & Scheuring (1926) in the lamellibranchs *Anodonta cygnaea* and *Unio pictorum*, where the algae lie in clusters in the posterior region of the mantle, especially near the siphons, and extend only as far as light penetrates, and by Boycott (1926) in *Limnaea peregra*, where cysts, full of *Chlorella*, appear in the foot, mantle edges, tentacles and other exposed tissues. But in these cases infection is rare and the algae invariably occur between the cells and not in them as in all cases of true association, including those described by Berkeley. Bidder (1920) has described intercellular association of *Syncrypta spongiarum* (Chrysomonadineae) with sponges, but the nature of this association remains obscure.

With the exception of the few cases mentioned above, including those of parasitic infection of ectodermal tissues, algae are always associated with animals which digest intracellularly (see Yonge, 1937*a*). Such animals are continually ingesting particles of the size of the algae, which seldom much exceed 10 μ in diameter, and it appears probable that the association originated by the permanent establishment within the tissues (or cell body in the case of Protozoa) of algae so ingested. If so, it must

have been established in different ways according to whether the animal was herbivorous or omnivorous (e.g. Protozoa, Porifera) or purely carnivorous (e.g. Coelenterata, Turbellaria). Conditions in the more highly evolved Mollusca vary widely and will be dealt with separately. Little is known about algal association in the Rotifera (see Buchner), but digestion here is certainly intracellular.

1. *Herbivores*

In a herbivorous animal the establishment of an association must involve specialization of the algae to resist digestion within the cells of the animal. The first stage in this is indicated in the Spongillidae (van Trigt, 1919). Both *Spongilla lacustris* and *Ephydatia fluviatilis* may occur in a green form due to the presence of a species of *Pleurococcus*. When isolated these algae lived for six months and longer, and multiplied; they normally occur free in fresh water. Colourless sponges were infected by them experimentally. Within the animal the algae increase, but sooner or later pass into a colourless state which is irreversible and precedes death. When a green sponge is placed in darkness the green algae diminish in number and the colourless forms increase. The algae are ingested by the choanocytes and passed to amoebocytes where they are always eventually digested, more quickly when other food is scarce. The green colour of the animal is maintained by the increase of those algae not digested and by the constant intake of new ones. The algae can do no more than prolong life under conditions in which other algae would be quickly digested. In the ciliate *Frontonia leucas* the zoochlorellae only remain in the animal so long as this is living in a stagnant and putrefying medium (Hood, 1927).

The conditions of association between *Paramecium bursaria* and *Chlorella*, described most recently by Parker (1926), Pringsheim (1928) and Loefer (1936*a*, *b*, 1938), indicates a further stage. The animal normally feeds on bacteria, algae and, especially, yeasts. The contained *Chlorella* are only consumed when the animal is starved and the animal does not long survive them. The association is a close one and only with difficulty could Pringsheim induce the animal to rid itself of the algae, most effectively by exposure to the combined effects of darkness, high temperature and lack of algal food salts, accompanied by good feeding of the animal. The white individuals so produced were easily reinfected by algae from other normal individuals; attempts with other races of *Chlorella* and other algae failed except partially with certain Ulotrichaceae. With cultures maintained in the light in a neutral medium with adequate supplies of nitrogen, as nitrate or ammonia, and of calcium, Pringsheim claimed that an autotrophic condition could be established in which both partners flourished in the absence of particulate food (other than bacteria). Loefer (1936*a*) has since succeeded in culturing *Paramecium bursaria* under bacteria-free conditions, but not indefinitely, which indicates that Pringsheim's results may have been due to the presence of bacteria in his cultures. He has also (1936*b*) been able to culture on agar the algae obtained from ageing *P. bursaria* cultures in which the animals gradually die off, leaving the plants dominant. He finds that these resemble *Chlorella ellipsoidea*. Here, therefore, both organisms can live independently, as in the Spongillidae, but normally do not do so, while only under exceptional circumstances does the animal consume the algae.

A probably even closer association exists in the numerous Radiolaria which contain brown zooxanthellae, the colonial forms in particular according to Brandt (see Buchner) only feeding holozoically when young and the algal content is low. But there is an unfortunate absence of modern work on association of algae with radiolarians. A well-developed association with zooxanthella also exists in many Foraminifera, e.g. *Orbitolites* (Doyle & Doyle, 1940).

2. *Carnivores*

In the more numerous cases of algal association with carnivores its effects are more far-reaching. Establishment of the association here demands specialization less of the plant than of the animal which normally (e.g. coelenterates) refuses either to pass plant material into the gut or to ingest it should it be taken in with the normal animal food. *Hydra* is incapable of digesting starch (Beutler, 1924). The work of Goetsch (1924) on association between members of the Hydrida and *Chlorella* indicates how this may have arisen. Neither *Pelmatohydra oligactis* nor *Hydra circumcincta* contain algae; indeed, the latter avoids light. In one case only was Goetsch able to infect the former by feeding it with tissues from *Anodon* parasitically infected with *Chlorella*. The algae were retained for over a week in the endoderm cells without harm to either party. Goetsch grafted the oral end of a green *Hydra viridescens* on to the aboral end of *Pelmatohydra oligactis* and found that the algal cells gradually disappeared from the oral end as the influence of the cells of *Pelmatohydra* extended; finally all disappeared. In brief, the cells of *Pelmatohydra* are not specialized for algal association. In *Hydra vulgaris* spontaneous infection was occasionally found, but experimental infection was easier and the animals could be induced to tolerate the algae permanently although the latter were always confined to certain regions of the endoderm where optimum conditions apparently prevailed. Artificial infection was easier still in *H. attenuata* (possibly, according to Goetsch, a variety of *H. vulgaris*); algae established themselves in all regions of the endoderm. In both these species infection only occurred when the animals were previously enfeebled, by lack of food, high temperature, etc., and it was accompanied by pathological symptoms often leading to death, and indicating, in the opinion of

Goetsch, a state of parasitism by the plant. In *H. attenuata* he found that the longer the experimentally induced association lasted the more difficult it was to dislodge the algae, and he suggests that in the course of time the tissues become so altered owing to the presence of the algae that a new race is formed (*H. viridescens*) in which association is normal but is only transmitted when reproduction is asexual. Finally, in *Chlorohydra viridissima* association is invariable, the next generation being infected by way of the egg (Haffner, 1925), and the algae are extremely difficult to dislodge while the resultant white *Chlorohydra* can then even be infected with a different alga, *Oocystis*. The colourless individuals are less resistant than green ones and can withstand hunger for a shorter time. Very similar conditions to those in *Chlorohydra* are recorded by Haffner (1925) in the fresh-water turbellarian *Dalyellia viridis*. Here also the *Chlorella* occur in the endoderm and in wandering cells in the parenchyma and are passed on by the egg. The algae can be removed and the animal can live without them as in *Chlorohydra*.

Thus it would appear that in carnivores the association takes place by way of an initial parasitism, postulated by Goetsch, followed by the establishment of tolerance by the animal, and finally, in *Chlorohydra* and *Dalyellia*, by that of a balanced, i.e. truly symbiotic, relationship, both members being capable of independent existence, although in nature the animals are probably never devoid of algae. After the establishment of such a balanced relationship further development appears to have taken place with one of three results: (*a*) by the animal becoming dependent on the plant, which remains capable of free existence, (*b*) by the plant becoming dependent on the animal and specialized for life exclusively within its tissues, and (*c*) by plant and animal becoming mutually interdependent, neither being capable of life apart from the other. Instances of all three types of association are given below.

(*a*) *Animal dependent on plant.* The best instance is the classic case of the marine turbellarian *Convoluta roscoffensis* (Gamble & Keeble, 1904; Keeble & Gamble, 1907; Keeble, 1910). Here the association is so intimate that, as these authors note, the algae really constitute an organ of the animal, and the two life histories are so intimately interwoven as to constitute a single individual. But whereas the animal must possess algae, and will not even develop properly in their absence, only individual members of the algae (probably a species of *Carteria* or *Chlamydomonas*) live in *Convoluta*, the great majority being free-living. It is unnecessary to go into the details of this well-known life history. The algae are passed on by the way of the egg case, the young being infected as soon as they hatch. In early life nutrition is holozoic, but later the organs of feeding degenerate and the animal lives on the excess material produced by the algae and then consumes these, gradually losing its green colour, and dying when this is completed but not before one or two batches of eggs have been laid.

An intermediate stage between *Dalyellia* and *Convoluta roscoffensis* is represented by *C. paradoxa* (Keeble, 1908). Here the associated organisms are brown zooxanthellae, although distinct from those found in coelenterates; they probably occur freely in the sea because uninfected larvae can be infected by placing them on sea weed from the zone normally occupied by the animals. This must be the normal mode of infection because the eggs are free from algae. The animal fails to develop in their absence, and to this extent is just as dependent on them as *C. roscoffensis*, but it never loses its alimentary organs and only digests the algae when starved. But when all are destroyed it can be reinfected, growth being then resumed. As in *C. roscoffensis* organs of excretion are absent, but the significance of this will be discussed later.

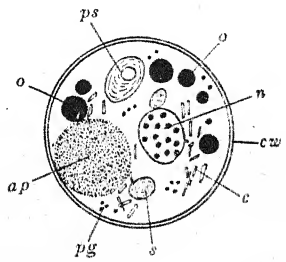


Fig. 1. Structure of a typical zooxanthella. Composite drawing. ×3000. *ap*, assimilation product; *c*, calcium oxalate crystals; *cw*, cell wall; *n*, nucleus; *o*, oil; *pg*, pigment granules; *ps*, primary starch granule; *s*, secondary starch granule. (After Doyle & Doyle, 1940.)

(*b*) *Plant dependent on animal.* Dependence by the algae on life in animal tissues is shown by the zooxanthellae which are associated with great numbers of shallow-water and surface-living animals in tropical seas. They invariably occur in all reef-building Madreporaria as well as in Actiniaria, Octocorallia, Scyphozoa such as *Cassiopea*, Hydrozoa (notably in *Millepora*), many compound ascidians, the Tridacnidae and many Foraminifera. In north temperate waters they are largely confined to a few Actiniaria, *Anemonia sulcata* being the commonest. These algae never occur free in the sea, they always increase by fission; no indication of sexual stages or spore formation has ever been seen. In the Metazoa the next generation is infected by way of the ovum or early embryo. Older accounts of these algae have largely been superseded by the recent, more detailed, work of Doyle & Doyle (1940) on the zooxanthellae from the foraminiferan *Orbitolites*, and also from certain Madreporaria, Siphonophora and Protozoa. The algae (Fig. 1) are always spherical, varying in diameter between 6 and 14 μ. A cellulose wall (*cw*) is well developed in some cases, especially in zooxanthellae from Madreporaria, but difficult to

determine in others, e.g. those from the Tridacnidae (Yonge, 1936). Doyle & Doyle consider that it is thinnest when growth is most rapid, and this may explain its poor development in the Tridacnidae where conditions for growth and multiplication are certainly ideal. Internally there is a large granular nucleus (*n*) difficult to see in life, a very prominent spherical assimilation product (*ap*) clearly visible in life, and a reticulate plastid with a colourless pyrenoid embedded in it, the whole being normally surrounded by amyloid material and forming the primary starch granule (*ps*). There may also be secondary starch grains (*s*), oil droplets (*o*) which are sometimes red in colour, pigment granules (*pg*) and crystals of calcium oxalate (*c*). Doyle & Doyle were able to show that the amount of starch varies directly with the intensity of the light to which the algae are normally exposed, those from the summit of coral heads having more starch than those from the under side. This again may account for the much greater quantities of starch noted by Yonge (1936) in the zooxanthellae from the Tridacnidae; these are always well exposed to light which is actually focused deep into the tissues (the mechanism for this is discussed later). In corals generally there is a greater accumulation of oil than starch in the algae. The concentration of calcium oxalate crystals was always found by Doyle & Doyle to be inversely proportional to the amount of starch, being thus most abundant in shaded areas. Experiments using *Orbitolites* showed that these crystals disappeared with accompanying increase in starch production when the carbon dioxide tension was increased in the presence of light; increase in oxygen or hydrogen tensions under similar conditions did not affect formation of starch or of calcium oxalate. In all cases starch diminished and calcium oxalate increased in the dark.* The probable nature of the association between zooxanthellae and the animals in which they live is discussed later.

The zooxanthellae in the Radiolaria are certainly different from those described above. They apparently do not occur in young individuals, each of which must therefore be infected afresh, probably by motile algal stages liberated when old Radiolaria die. For full details see the papers of Brandt and others quoted by Buchner (1930). The systematic position of these and other 'symbiotic' algae is discussed by Fritsch (1935).

(*c*) *Interdependence of plant and animal.* The one group in which the animal appears to be as dependent on the zooxanthellae as they are upon the animal is the Alcyonacea. Ashworth (1899) and Pratt (1906) originally showed that the digestive organs in certain genera are reduced or absent (details are given later), while Gohar (1940) has recently produced experimental evidence that the Xeniidae are incapable of capturing animal prey and quickly show signs of starvation when kept in the dark although animal food is present. On the other hand, animals starved in the light retain health and vigour, presumably relying on the zooxanthellae for nutrition.

* I have discussed this matter with Prof. M. Skene, who informs me that the role of calcium oxalate in plant metabolism remains obscure.

3. *Infection by previously specialized algae*

In addition to the above categories there are other instances of association with algae which can most easily be explained on the assumption (indeed certainty in some cases) that the animal has been infected by algae already specialized for life within another type of animal. This is certainly the case in the nudibranch *Aeolidiella* (*Eolidina*) *alderi* (Naville, 1926; Graham, 1938), which feeds exclusively on *Heliactis bellis*, one of the few temperate-water Actiniaria which contain zooxanthellae. These are ingested as if they were portions of food, but cannot be digested, and they remain in the endoderm cells of the cerata along with the nematocysts which are acquired in the same manner. Naville believes that they increase in the nudibranch, but Graham has been unable to confirm this. It is questionable whether this can be regarded as more than an incidental case of association, but it may explain the mode of origin of the more intimate association in other opisthobranchs, e.g. *Aeolis glauca, Doridoeides gardineri, Melibe rangii, Phyllirhoë, Spurilla neapolitana* and *Favorinus albus* (see Buchner, 1930, for further details and references). In *Tridachia crispata*, an opisthobranch of the order Ascoglossa, zooxanthellae are always present in a restricted zone a short distance from the margin of the undulating body folds (Yonge & Nicholas, 1940). They occur freely within the connective tissue and increase by division. The animal is a highly specialized herbivore. It is very unlikely, in view of the specialized nature of zooxanthellae, that infection was originally parasitic as in the case, already recorded, of *Chlorella*, which occasionally infects fresh-water Mollusca. It seems inherently more probable that the zooxanthellae had already been specialized for life within a carnivorous animal and so were able to resist digestion when originally ingested by *Tridachia*.

The same mode of origin must be postulated in the case of the association with zooxanthellae found in the only family of Lamellibranchia in which this occurs, the Tridacnidae (Yonge, 1936). Like all lamellibranchs except the Septibranchia these animals are herbivorous, feeding on phytoplankton. Yet, as will be shown later, the association here is more complex and has more far-reaching effects on the structure and habits of the animal than even in *Convoluta roscoffensis*. Finally, there is the case of compound ascidians already mentioned where the zooxanthellae must almost certainly have been acquired after specialization elsewhere for this type of life.

II. MODIFICATIONS RESULTING FROM ASSOCIATION

As a result of their association both plants and animals may be modified in structure and function, and also in habits. These modifications, moreover, may affect either the individual or the race.

1. *Algae*

Here structural changes affecting the individual are exemplified by the smaller pyrenoids claimed by Haffner (1925) to occur in *Chlorella* living in *Chlorohydra* as compared with free-living individuals (these may, however, be different races), and, more obviously, by the effects of life in *Convoluta roscoffensis* on the contained algae (Keeble & Gamble, 1907). When free-living these possess four equal flagella, frequently a stigma, a conspicuous nucleus and pyrenoid and a cell wall; they are capable of saprophytic existence when they may be colourless. It is in the latter stage that they are normally attracted chemotactically to the egg case of *C. roscoffensis*, and, after vegetative increase, infect the newly hatched animal. There they develop chlorophyll, although ingestion of active or temporarily resting green individuals may also cause infection. Within the tissues the algae divide, flagella never appearing; although there are transient stigmata, there are no cell walls, while the nucleus undergoes progressive degeneration and finally disappears. Only the mechanism of photosynthesis remains unaffected. Physiological changes are exemplified by the power of saprophytic nutrition acquired by the *Chlorella* of *Chlorohydra* (Haffner, 1925) and *Paramecium bursaria* (Pringsheim, 1928). The race is affected in the case of those zooxanthellae which have lost the capacity for free existence and sexual reproduction and have acquired a relatively thick cellulose wall.

2. *Structure of animals*

The best instance of the influence of the association on the structure of the individual is provided by the gradual change from *Hydra attenuata* to *H. viridescens* claimed by Goetsch (1924) to occur as a result of infection by *Chlorella*. Changes affecting the race are numerous, the best examples being provided by *Convoluta paradoxa* and *C. roscoffensis*, by certain Alcyonacea and by the Tridacnidae.

In both species of *Convoluta* excretory organs are absent, while in *C. roscoffensis* the effects on the bodily structure are profound owing to the eventual loss of the alimentary system. The unusually omnivorous habit, compared with the strictly carnivorous diet of most Turbellaria, may also be a consequence of its acquired power of feeding on its contained algae. In tropical Alcyonacea, Pratt (1906) has shown a progressive reduction in the ventral and lateral mesenterial filaments (i.e. the regions concerned with secretion of the extracellular protease essential for the preliminary digestion of animal food in all Anthozoa) in correlation with the number of zooxanthellae present. In *Lobophytum* conditions approximate to those in *Alcyonium* of temperate waters with little reduction in the filaments and few algae. In *Sarcophytum* the filaments are smaller with fewer gland cells, algae being more numerous. In tropical species of *Alcyonium* the filaments are extremely small with few gland cells while zooxanthellae are very numerous. In *Sclerophytum* the filaments are very small or absent and hence gland cells are very few. The algae are extremely numerous. Moreover, in this genus the polyps are few in number and the tentacles minute. In *Xenia* and *Heteroxenia*, Ashworth (1899) had previously shown that only the dorsal (exclusively current-forming) filaments are present. Ashworth thought that the digestive function might have been transferred to the stomodaeum in which he found numerous goblet cells, but Gohar

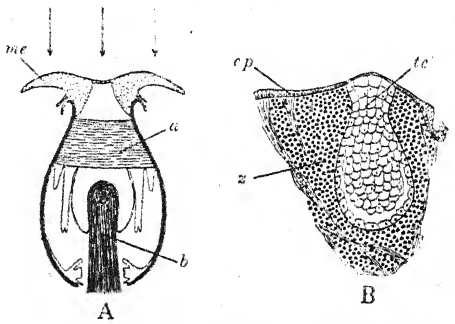


Fig. 2. *Tridacna crocea*. A, diagrammatic cross-section through region of adductor muscle (*a*) and byssus (*b*), showing enlarged and fully exposed mantle edges (*me*) containing zooxanthellae. Arrows indicate direction of light rays. B, section through mantle edge showing hyaline organ. ×72. *ep*, epithelium of mantle edge; *tc*, transparent cells of hyaline organ; *z*, zooxanthellae in blood spaces around hyaline organ. (A, original; B, after Yonge, 1936.)

(1940) has been unable to confirm the presence of these, while he points out that the stomodaeum is actually unusually short and that little digestive action could take place during the passage of food through this region. The tentacles are not reduced; indeed, unlike those of *Sclerophytum*, they are very large, but they do not react to animal food. Gohar states that small animals 'promenaded on the tentacles and on the peristomial disks'; on the other hand, the mouth is very small, and he never found food within the coelenteron. The results of his experiments with animals kept in light and darkness have already been referred to. Conditions in the Xeniidae and in *Sclerophytum* would, therefore, appear to approximate to those in *Convoluta roscoffensis*.

The effect of algal association on the structure of the animal is greatest in the Tridacnidae (Yonge, 1936). Here the zooxanthellae are housed in vast numbers within the greatly thickened and laterally extended mantle edge (Fig. 2 A). Although the

animal is secured by a byssus (*b*) (or, in the larger species, purely by its own weight in adult life), the mantle edges (*me*) with their contained algae are topographically uppermost, i.e. the mantle and shell have rotated, clockwise when viewed from the right side, to an angle of almost 180° in relation to the visceral mass and foot. This movement, indicated in Fig. 3, has brought the umbo (*u*) and hinge close to the foot, while the mantle edges have moved round from the ventral to the dorsal side of the visceral mass. The unique features of the morphology of the Tridacnidae have long been known, but the view that the presence of the algae provide the cause for this has only recently been advanced (Yonge, 1936). The mantle edges are invariably widely exposed to the light (Fig. 2 A), even in the full heat of the midday tropical sun. The animals are confined to shallow water, usually near the summits of coral reefs. The light is focused deep into the tissues by the aid of 'hyaline organs', in the form of lenses. These were formerly described as eyes, but they have no nerve and no retina, while around them the zooxanthellae (*z*) congregate in maximum numbers (see Fig. 2 B). The algae are invariably contained within phagocytic blood cells and increase by division. They are also found in all stages of digestion in similar cells in the region around the digestive diverticula which are somewhat reduced in bulk compared with other lamellibranchs. Apart from this slight reduction the organs of feeding and digestion are normal and the animal certainly also feeds on phytoplankton like other lamellibranchs. But it is possible that the extra food represented by the algae which it farms in its mantle tissues may account for the immense size attained by the giant clam, *Tridacna deresa*, by enabling it to exceed the limits of size imposed by ciliary feeding. A remaining peculiarity is the abnormal size of the kidneys (Fig. 3 B, *k*), fused in the middle line, which is probably due to the need for getting rid of the relatively immense quantities of indigestible material carried to them by the phagocytic blood cells after digestion of the zooxanthellae. What corresponds to faecal material has to be voided by way of the kidneys, where much of it accumulates in the form of spherical concretions.

3. *Habits of animals*

The habits of all animals normally associated with algae are influenced by the need for keeping in well-illuminated water (except in *Frontonia* (Hood, 1927), where a stagnant, putrifying medium is essential). Where the animal is sedentary, e.g. Anthozoa or Tridacnidae, this probably only affects the larvae which must settle where the adult will be in the light. Kawaguti (1941) has pointed out that the range of light for positive phototaxis varies for the planulae of different species of reef-building corals and appears to be associated with varying content of zooxanthellae. Abe (1939) has suggested that expansion of reef-building corals by night only may be due to a nightly accumulation of carbon dioxide when photosynthesis by the algae is impossible. It is difficult to accept this contention which is discussed elsewhere (Yonge, 1940) together with data indicating the great influence of light on the growth and distribution of reef-building corals. It is uncertain how much of this influence is due to the presence of zooxanthellae, but some of it certainly must be. Species of the semi-sedentary scyphozoan *Cassiopea*, which live in shallow tropical waters with the subumbrella surface uppermost, contain zooxanthellae

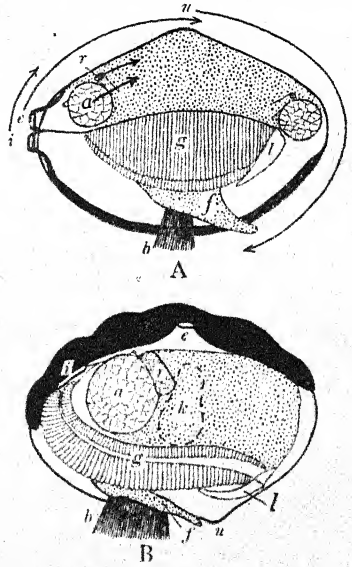


Fig. 3. Diagrams indicating the manner in which the mantle, with the shell, has rotated in relation to the visceral organs and foot in the Tridacnidae, the result being to bring the mantle edges (shown black) topographically uppermost. A, lateral view of unspecialized, dimyarian lamellibranch. Arrows indicate the rotation of the umbo (*u*), exhalant and inhalant siphons (*e*, *i*), posterior adductor (*a*) and pedal retractor (*r*) involved in the evolution of the Tridacnidae. B, lateral view of *Tridacna crocea* showing result of rotation which involves disappearance of the anterior adductor. The gills (*g*), labial palps (*p*), foot (*f*) and byssus (*b*) remain in the same relative positions. The extent of the large kidney (*k*) is indicated by the broken lines.

(Bigelow, 1900; Smith, 1936) which may have some influence on their habits. The habits of zooplankton containing zooxanthellae are affected; they tend to remain in the surface waters during midday when other forms migrate into deep water (Russell, 1927). This is true of Radiolaria (Haecker, 1908), the zoantharian larvae *Zoanthella* and *Zoanthina* (Conklin, 1908), and the ephyrae of *Mastigias papua* (Uchida, 1926). Pascher (1930) has shown that in *Amoeba stigmatica* the presence of zoochlorellae is associated with that of a large stigma and pronounced phototaxis.

But the best instance of effect on the habits of the animal is again shown by *Convoluta roscoffensis* (Gamble & Keeble, 1904). This species lives in sand at the upper limit of high-water mark of lowest neap tides and always within an area of drainage, i.e. where it will get maximum light exposure combined with safety from desiccation. Positive phototaxis causes migration to the surface (which never occurs during low tide at night), while positive geotaxis induced by vibrations (easily reproduced by tapping the ground lightly around an exposed colony) causes migration downward when the tide reaches them. There are also fortnightly lunar variations due to periodicity of reproduction; the animals lay their egg capsules below the surface at the beginning of spring tides and the body usually ruptures, the hinder half being left in the sand and only the head end rises. As a result the patches are reduced at neap tides. *C. paradoxa*, on the other hand, lives on weed about low-water mark and migrates horizontally with the spring and neap tides, thus obtaining maximum illumination without danger of desiccation (Keeble, 1908).

III. SIGNIFICANCE OF THE ASSOCIATION

The significance of the association varies widely. Only in a few instances can it be regarded as a balanced relationship or true symbiosis, for which reason that term has not been used. It will be simpler to discuss the matter first in terms of the algae and then of the animals.

1. *Algae*

The plants certainly obtain protection and food. Life within an animal is more secure than in the surrounding medium unless the alga is eventually digested, as in *Spongilla*. Food is more important. This consists of carbon dioxide and inorganic salts, of which those containing available nitrogen (as nitrates or ammonia) or phosphorus (as phosphates) are the most important. The utilization of carbon dioxide by the *Chlorella* in *Paramecium bursaria* and by the zooxanthellae in various reef-building corals has been demonstrated experimentally by Parker (1926) and by Yonge & Nicholls (1931*a*), but has probably little significance in any association because carbon dioxide is normally abundant in both the sea and fresh water. It is otherwise with nitrogen and phosphorus which are limiting factors for plant growth. Pütter (1911), working on the actinian *Aiptasia* which has zooxanthellae, found that if the surrounding water contained no ammonia, small quantities were excreted into it by the animal; if the water contained a total of between 0·113 and 0·266 mg. of ammonia, no appreciable change took place, but if the ammonia content exceeded this amount, then there was a reduction of about one-half in the course of the experiment. He concluded that the zooxanthellae were responsible. In *Convoluta roscoffensis* Keeble & Gamble (1907) found that the green cells utilized nitrogen better in the form of uric acid or urea than of nitrates.

Data on phosphorus utilization are more abundant. Yonge & Nicholls (1931*a*) found that the phosphorus content of normal sea water (containing 3·41 mg./cu. m.) was either reduced to zero or only slightly increased by species of many genera of reef-building corals in experiments for 4 days in jars containing 2500 c.c. of water. In parallel experiments with *Dendrophyllia*, a deep- and cold-water coral which has extended its vertical range in the tropics but never contains zooxanthellae, the phosphorus content increased greatly, in one case by as much as 3400%. When the phosphorus content was increased to 2036 mg./cu. m. the results were still more arresting, as shown in Fig. 4. With *Psammocora*

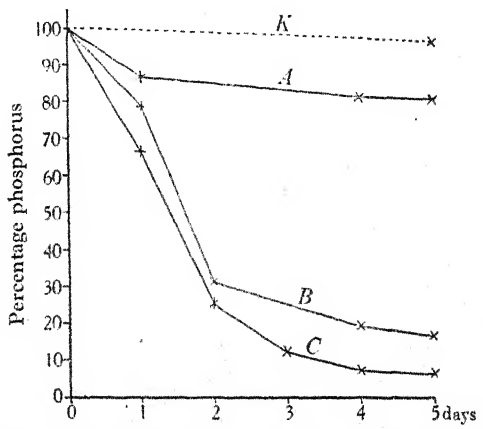


Fig. 4. Graphs showing percentage change in phosphorus content (original value 2036 mg./cu. m.) of sea water in jars of 2500 c.c. capacity containing corals. A, *Psammocora*; B, *Porites*; C, *Favia*; K, control. (From Yonge & Nicholls, 1931*a*.)

(A) the phosphorus content dropped to about 81% after 5 days, with *Porites* (B) to just under 20% and with *Favia* (C) to about 7% of the original very high total. In the control (K) the phosphorus dropped only to 96·6%. This indicates that the zooxanthellae are capable not only of removing all the phosphorus excreted by the animal but can utilize it in great amounts from the surrounding water, far more than would be present normally. Later experiments carried out with corals kept in darkness for 152 days and largely deprived of their zooxanthellae invariably showed percentage *increases* in the phosphorus content, in some by as much as 688% in one day. All of this must normally be removed at once by the zooxanthellae. In *Tridacna crocea*, Yonge (1936) found that all trace of phosphorus was removed from 2500 c.c. of normal sea water in 24 hr., whereas with specimens of another lamellibranch, *Spondylus*, the

phosphorus content was increased from 4 mg./cu. m. to 65, 109 and 486 mg. in the three animals used. Smith (1936) found a similar utilization of phosphorus from the surrounding water by the zooxanthellae in *Cassiopea*, and later (1939) that the phosphorus excretion in *Anemonia sulcata* (containing zooxanthellae) is about half that of *Actinia equina* (with no zooxanthellae).

In addition, the algae may obtain food saprophytically from the animal. The *Chlorella* in *Paramecium bursaria* (Parker, 1926; Pringsheim, 1928) and in *Chlorohydra* (Haffner, 1925) will survive and even increase in the *dark* if the animals are well fed, which indicates that nutrition must be saprophytic. The utilization of uric acid and urea in *Convoluta roscoffensis* referred to above also indicates saprophytism.

It may be concluded that algae in association with animals are at an advantage over those living freely in their ability to tap at the source valuable supplies of nitrogen and phosphorus.

2. *Animals*

Three possible advantages to the animal have been postulated, namely, additional supplies of oxygen from the photosynthetic activities of the plants, nutriment in one form or another from these, and the automatic removal by them of the waste products of metabolism.

(i) *Oxygen*

There can be no doubt about the relatively large amounts of oxygen produced by the algae. References to earlier literature are given by Yonge, Yonge & Nicholls (1932) in a paper which gives considerable data for conditions in reef-building corals. These, when exposed for 9 hr. in the light, almost invariably produced more oxygen than the animal (and zooxanthellae) utilized over that period; at a depth of 4 m., however, the oxygen production over 24 hr. was, with one exception, lower than consumption. About midsummer the peak in oxygen production at the surface occurred between 10 and 11 a.m., the later falling off in photosynthesis being probably due to the accumulation of end-products. Reef-building corals survived for 2 weeks and more in sealed jars with, in some cases, an actual increase in oxygen content at the end of this period, but the *p*H always fell, indicating that more carbon dioxide is produced than the algae can utilize. Further data have been provided by Marshall (1932) for coral planulae, by Kawaguti (1937) for other Madreporaria, by Welsh (1936) for a marine turbellarian, *Amphiscolops langerhansii*, and by Smith (1939) for *Anemonia sulcata*.

A number of authors have stressed the possible importance in the life of the animal of the oxygen so produced, pointing out that fresh-water animals with algae, e.g. *Spongilla* and *Chlorohydra* (van Trigt, 1919; Haffner, 1925), live longer in deoxygenated water than do colourless ones. But the significance of this in nature is probably slight, except in green *Frontonia* which live only in stagnant water (Hood, 1927). Smith (1939) thinks that *Anemonia* can better survive conditions in poorly oxygenated pools when the tide is out. But the vast majority of temperate shore actinians possess no zooxanthellae and yet flourish in rock pools, while in the Coelenterata respiration is largely unaffected by oxygen tension (see Yonge *et al.* (1932) and earlier work there quoted). A more fundamental question is raised by Verwey (1931). On the basis of experiments carried out with *Acropora hebes* he calculated that an ordinary colony of a large *Acropora*, weighing several kilograms, would consume during one night 250 c.c. of oxygen for every kilogram weight, and hence a reef of thousands of kilograms would consume hundreds of litres of oxygen. He maintained that the oxygen produced by the zooxanthellae during the daytime is essential for the respiratory needs of the coral at night. The validity of his conclusions depends on that of the original figures for respiration. Mayor (1918, 1924) estimated the oxygen consumption for a variety of corals, relating this to the amount of living tissue. He obtained widely different results, *Acropora muricata* for instance having apparently a respiratory rate per unit of living tissue more than 18 times that of *Siderastrea radians*, other species coming in between. As Verwey (1931) points out, the former figures are equivalent to those for an active animal such as a fish or a squid. Clearly these widely divergent results for closely related species of corals demand an explanation. Yonge (1937*b*) has shown that much of the *apparent* utilization of oxygen by corals is due to the oxidation of the mucus secreted by them during the course of the experiment (or by respiration of bacteria present in the mucus). The production of mucus varies greatly in different genera and is exceptionally high in *Acropora* which always literally drips with it when handled. Secretion in *Siderastrea*, on the other hand, is very low. It is thus impossible to accept at their face value figures which claim to represent either absolute or relative rates of respiration in corals, or conclusions based on such data. It is very unfortunate that Verwey should have based his conclusions on figures for the apparent rate of respiration in a species of *Acropora* which, judging from Mayor's results, may be as much as 18 times too high; even if only 3 times too high they invalidate Verwey's thesis.

In general it may be said that, although it is possible that under certain circumstances the animal may benefit from the oxygen liberated during photosynthesis by the algae, the importance of this remains uncertain.

(ii) *Nutrition*

There is sound evidence that in many, though not all, cases the animals do obtain food from the algae. This may occur in one of three ways.

(*a*) Organic matter may pass from the living algae into the tissues of the animal. This occurs according

to Pringsheim (1928) in *Paramecium bursaria*, while Goetsch (1924), though not Haffner (1925), believes it is true for *Chlorohydra*. Gamble & Keeble (1904) have shown that this occurs throughout life in *Convoluta paradoxa* and in early life in *C. roscoffensis*. The opinions of earlier workers on this matter are summarized by Buchner (1930). It should be noted that this process implies excessive production by the algae, presumably owing to the relatively small numbers present compared with the food available in the animal. Such conditions do not prevail in the Madreporaria, where, as already noted, additional nutrient salts are taken from the surrounding water.

(*b*) Dead algae may be digested by the animal. This, according to Haffner (1925), is what occurs in *Chlorohydra* and *Dalyellia*. Doyle & Doyle (1940), who found abundant starch grains, calcium oxalate crystals and pigmented oil droplets in the cytoplasm of *Orbitolites* only when the contained zooxanthellae were distended with the same substances, suggest that transference into the foraminiferan occurs when such 'overstuffed' zooxanthellae disintegrate during division.

(*c*) Algae are killed and digested by the animal in a number of cases; always in the Spongillidae (van Trigt, 1919), in the latter stages of life in *Convoluta roscoffensis* (Keeble & Gamble, 1907), in the Xeniidae (Gohar, 1940) and in the Tridacnidae (Yonge, 1936). The stimulus which causes the blood cells in the Tridacnidae to move from the mantle edges to the region of the digestive diverticula there to digest the zooxanthellae which have previously flourished within them remains obscure. In extreme cases of starvation the algae in *Paramecium bursaria* (Pringsheim, 1928) and in *Convoluta paradoxa* (Keeble, 1908) are also killed and digested.

Many workers (see Buchner, 1930) have shown that 'green' animals will survive starvation for longer periods than 'white' individuals of the same species, but details of the exact manner in which nutrition is obtained are lacking. The case of the Madreporaria and other coelenterates which harbour zooxanthellae (with the exception of the Xeniidae and other Alcyonacea) is different. There has been extensive controversy about the part played by the zooxanthellae in the life of these animals. One group of workers believe that they form an important part of the food of corals, another that they do not. The evidence has been reviewed by Boschma (1924, 1925*a–c*, 1926, 1929) and Yonge & Nicholls (1931*b*). Boschma held the former view largely as a result of his discovery that during starvation degenerating algae accumulate in the digestive zone of the mesenteries, at the base of the filaments (see Fig. 5). It was later shown independently by Mouchet (1930) and Yonge (1931*a*) that in Actiniaria and Madreporaria respectively this digestive zone is also the excretory zone, indeed, the only region where particulate matter passes into or out of the bodies of these animals. Smith (1936) has found that the same is true of the gastric filaments in the scyphozoan *Cassiopea*.

It has been shown by Yonge & Nicholls (1931*a*, *b*) that whenever the metabolism of corals is lowered, by starvation or by exposure to high temperature or low oxygen tension, the zooxanthellae are expelled (Fig. 5, *z*), though often apparently quite healthy; the same occurs when corals, either fed or starved, are kept in darkness. In the former instances the algae were starved of nutrient salts, in the latter of the means of synthesizing carbohydrate. In all cases the zooxanthellae were carried to the excretory zone in wandering cells (Fig. 5) and eventually passed out of the mouth in the form of dark, mucus-laden masses. The presence of degenerating algae in the 'digestive' zone of the mesenteries does *not*, therefore, indicate that they are being digested. Carefully controlled experiments in which a wide variety of reef-building corals were fed and starved

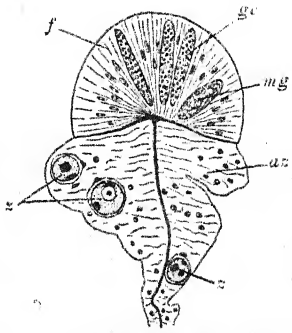


Fig. 5. Transverse section through the end of a mesentery and the mesenterial filament of the madreporarian *Psammocora gonagra* after starvation in light for 166 days. ×720.' *az*, absorptive (and excretory) zone of mesentery; *f*, mesenterial filament with contained gland cells (*gc*) secreting protease and mucous gland (*mg*); *z*, degenerating zooxanthellae in process of being extruded. (After Yonge & Nicholls, 1931*b*.)

in light and in darkness revealed that when starved in light expulsion of zooxanthellae began almost immediately, and shortly afterwards the tissues of the coral began to diminish—although abundant zooxanthellae remained for many months in the tissues. Retreat of the tissues was most marked in species of *Fungia* owing to the configuration of this genus; at the end of 73 days in one instance the greater part of the upper surface of the skeleton was exposed, the greatest diameter of the 'mouth' opening being 4·5 cm., that of the animal being 6·5 cm., yet abundant zooxanthellae remained although they were continually being extruded. Similar effects were noted in many other genera, notably in *Lobophyllia*, where the edge-zone retreated to a marked extent. The onset of the effects of starvation in corals with zooxanthellae always occurred as quickly as in *Dendrophyllia* where algae are absent. Loss of zooxanthellae in the dark occurred at similar speeds in

both starved and fed animals, not more quickly in the former. This was indicated by estimations of the oxygen exchange in light and darkness of four specimens of *Psammocora gonagra*, two of which had been fed and two starved for the previous 137 days, one of each in light and the others in darkness. The difference between the figures for oxygen exchange in light and darkness for each specimen revealed the amount of oxygen produced by the contained zooxanthellae in the light and so provided an indication of their number. Compared with the control (fed in light) the specimen starved in light retained 64% of the algae originally present, while both specimens kept in darkness showed a reduction to about 40%. The first reduction, to 64%, may be attributed to the effect on the algae of the lowered amounts of nitrogen and phosphorus excreted by the starved animal, the further reduction to 40% in the other two specimens being due to the added effect on carbohydrate synthesis of the absence of light. But it is notable that the specimen starved in darkness showed no evidence of further reduction in algal content owing to their digestion by the starved animal. These results do indicate that Madreporaria under no circumstances digest their contained zooxanthellae, while experiments on *Fungia* give no support to the later suggestion by Gohar (1940) that the corals obtain some nutriment from algae which degenerate before or during expulsion.

Reef-building corals can live in dark places with few if any contained zooxanthellae both in nature and under experimental conditions (see Yonge & Nicholls, 1931*a*, for references). In Actiniaria, Fulton (1921) showed that *Actinia bermudensis*, which contains zooxanthellae, can survive unfavourable conditions as well in darkness as in light, although he believed, without positive evidence, that the actinian feeds on the algae. Smith (1939) has confirmed the earlier findings of Brandt (see Buchner, 1930) that *Anemonia sulcata* kept in darkness loses its zooxanthellae, but also found that these animals can live indefinitely under such conditions if well fed and aerated. There is thus positive evidence that in Madreporaria and Actiniaria association with zooxanthellae is *not* essential to the life of individual colonies or animals.

(iii) *Removal of waste products of metabolism*

It follows from what has already been stated about the utilization of carbon dioxide, nitrogen and phosphorus by the algae that these act as automatic organs of excretion to the animals. To this can be attributed the loss of the organs of excretion in both species of *Convoluta* (Keeble & Gamble, 1907; Keeble, 1908). On the other hand, because they have to dispose of the remnants of the zooxanthellae after digestion within the blood cells, the kidneys of the Tridacnidae are enlarged although all available phosphorus is removed by the algae (Yonge, 1936). The significance of this factor is more difficult to assess in Protozoa, Porifera and Coelenterata which possess no specialized organs of excretion, but it may well be significant owing to the increased efficiency in metabolism due to the automatic removal of end-products. In the case of the reef-building corals, although actual experimental proof is lacking, a previous opinion may be quoted, namely, that the association between corals and zooxanthellae is essential to the plants, certainly not to individual coral colonies, but, by assisting in the speed of metabolism, probably an indispensable factor in the necessarily exceptional powers of growth and repair possessed by the marine communities known as coral reefs (Yonge, 1931*b*). In this connexion, Thiel (1929) has suggested that both in the corals and in the Tridacnidae the production of oxygen in the tissues may assist in the formation of the skeleton, while Verwey (1931) points out that an excess of carbon dioxide in the tissues might cause dissolution of the skeleton. These views are worthy of experimental analysis, but it remains true that corals such as *Dendrophyllia* and other tropical lamellibranchs, such as *Chama* and *Spondylus*, form equally massive skeletons without the assistance of zooxanthellae.

IV. INTERRELATIONSHIP BETWEEN ASSOCIATES

A true symbiosis, that is, a balanced relationship, is certainly of great advantage to both parties, because an almost closed system is thus established requiring the minimum of energy, apart from light, from outside sources. This is probably true of *Paramecium bursaria*, especially when living autotrophically (actually probably obtaining bacteria), but is probably best displayed in *Chlorohydra* and *Dalyellia* and to some extent in the coelenterates (other than certain Alcyonacea) possessing zooxanthellae, although here the plants are incapable of free life. But in all cases the normal cycle of events is shortened, the plant taking inorganic nutriment from the animal which has this excrement automatically removed and, of more dubious value, obtaining additional oxygen, while in some cases organic matter is passed from the intact alga to the animal.

Two interesting cases of experimental association illustrate the above points. The first was made between *Chlorella* sp. and *Azotobacter chroococcum* (Lipman & Teakle, 1925), the other between *Chlorella pyrenoidosa* and embryonic connective tissue cells in chick embryonic extract and chicken plasma medium (Buchsbaum, 1937). In the former both associates grew better than when cultured alone, the bacteria fixing greater quantities of nitrogen which the algae were able to utilize. In the latter both parties grew better when associated but only if exposed to light. Buchsbaum attributes the advantage purely to gaseous exchange, the animal cells having excess carbon dioxide removed and acquiring additional oxygen while the algae gained carbon dioxide. In

the conditions of his small, sealed drop cultures this interpretation may well be correct, but the algae have also, of course, abundant sources of nitrogen. The better growth of the cultures in darkness without algae may well be due, as Buchsbaum suggests, to greater accumulation of carbon dioxide when algae are present, but this does not necessarily cut out the probability of the supply of nitrogen being an important factor in the light because in the dark, with the stoppage of photosynthesis, protein synthesis must also be curtailed. But this work is of particular interest as demonstrating experimentally the mutual advantage of intimate association between plants and animal tissues.

The association ceases to be truly symbiotic when one party exploits the other. The algae never exploit the animal except in the early, parasitic stages described by Goetsch. Unless a balanced relationship is attained the plant ends by destroying itself as well as the animal. The animal may, however, exploit the algae. This is done, although to the ultimate destruction of the individual, by *Convoluta roscoffensis* and, without destruction of the individual, in the Tridacnidae, *Orbitolites* and probably the Radiolaria and also, to a less extent, in *Convoluta paradoxa* and the Spongillidae. The position in the Xeniidae and similar Alyconacea is obscure because here the animal is apparently unable to feed yet apparently can exist indefinitely on its contained zooxanthellae; more information is required for the full analysis of this case. But in other cases certainly destruction of the individual is avoided because the animal never ceases to feed holozoically and obtain some energy from outside sources. For the full exploitation of the algae it is necessary for the animal to maintain these alive permanently in its tissues and yet be able, when necessary, to digest them. *Tridacna*, which has developed 'farming' of the algae to a greater extent than the allied *Hippopus*, is the supreme example of exploitation of associated algae by animals.

V. IMPRISONED PHYTOPLANKTON

Although association with unicellular algae occurs in isolated instances in fresh waters and in temperate seas, it is widespread in tropical seas where a large proportion of the shallow-water and shore invertebrates contain zooxanthellae. The number of algae is incalculable, but some idea may be gained from the statements of Marshall (1932) who found up to 7400 algae in a single planula of *Porites* sp. not more than 1 mm. long and 0·5 mm. wide and estimated a total of 25,000 in the somewhat larger planulae of *Pocillopora bulbosa*, and of Doyle & Doyle (1940) who found an average content of 16,000 zooxanthellae in *Orbitolithes duplex* some 2 mm. in diameter. The endodermal tissues of the vast expanse of Madreporaria which constitute the bulk of Indo-Pacific and Atlantic reefs are stocked in much the same concentration, as are the extensive fields of Alcyonacea in the Indo-Pacific and of Gorgonacea in the Atlantic. So far as the Madreporaria at least are concerned, the population of zooxanthellae appears always to be controlled not by the space available in the tissues but by the quantities of inorganic food salts. It has already been shown that the zooxanthellae will take up additional phosphorus from the surrounding water, while data on oxygen exchange between corals and zooxanthellae (Yonge *et al.* 1932) indicate that above a certain degree of illumination, represented by varying depths according to latitude, silt content, etc., the concentration of algae in the tissues remains approximately constant. The plants are apparently incapable of increasing in direct proportion to the increase in light. This is probably due to limitation imposed by the supply of inorganic food salts excreted by the corals, but it is also possible that the zooxanthellae function better when the light is not too intense. The zone of maximum coral growth is some little distance below the surface and this may possibly be correlated with the greater efficiency of the algae at these depths (see Yonge, 1940, for full discussion on this).

This immense bulk of plant life represents a form of 'imprisoned phytoplankton' (Yonge, 1931*b*), the significance of which cannot be overlooked in any account of conditions of life in shallow tropical waters. Some comparison with conditions of life in free phytoplankton may be made by reference to the recent review by Harvey (1942). Imprisoned phytoplankton live under more stable conditions than the free forms; only major changes influencing the animals with which they are associated will affect the algae. For instance, exposure to extreme light and heat on the surface of reefs at low tide in the summer may cause complete ejection of zooxanthellae by the corals (Yonge & Nicholls, 1931*a*), the normal population not being regained for some three months. There is a more constant supply of food owing to the steady metabolism of the animal, although there is probably some slight seasonal fluctuation. So far as Madreporaria, Actiniaria and Gorgonacea are concerned there is no 'grazing effect' by the animal such as is produced by zooplanktonic organisms on the free phytoplankton; but there is in the Tridacnidae, in the Xeniidae and some other Alcyonacea and in the Foraminifera. But the total quantity of imprisoned phytoplankton depends directly on that of the associated animals. The major effects of association are on the quantity of the free phytoplankton. No nutrient salts are passed into the water by any animal possessing associated algae; all are intercepted by the imprisoned plants. This means that there will be less free phytoplankton and hence less zooplankton. This in turn will have its effects on the original animals because this zooplankton is the normal food of Coelenterata. But at the same time the increased supply of zooplankton which would be available were there no imprisoned phytoplankton would be widely scattered and only a small propor-

tion of it would find its way within range of the tentacles of these sedentary animals. If the theory that the presence of associated algae *is* important to reef-builders owing to the increased metabolic efficienty it confers be correct, then it is probably better to have slightly less available zooplankton with automatic excretion than more zooplankton without this aid.

In connexion with their hypothesis that zooplankton are excluded from regions of high phytoplankton content, Hardy & Gunther (1935) suggest that tropical planktonic and shallow-water animals which contain associated algae may have been able to overcome this excluding influence by 'some counteracting physiological process'. They contrast polar seas with their high nutritive salt content, rich 'free' phytoplankton and exclusion effects with tropical seas with very low nutrient salt, 'imprisoned' phytoplankton and no exclusion effects. They suggest a possible correlation between the distribution of coral reefs and that of water masses which have passed furthest from regions of predominantly 'free' phytoplankton or between reefs and waters where the phytoplankton is never sufficient to produce exclusion effects. The basis of their exclusion theory cannot be discussed here but, whatever its validity with regard to zooplankton, there are certainly no data to justify the suggestion that reefs have been formed only because their principal components have overcome exclusion from surface waters owing to their 'imprisonment' of the phytoplankton. The presence of algae within the tissues of so many tropical marine invertebrates may, on the basis of existing knowledge, be attributed from the standpoint of the algae to the result of the intense competition in these waters for the limited supplies of nitrogen and phosphorus, and from that of the animals to greater survival value owing to the increased metabolic efficiency gained by some, e.g. Madreporaria, Actiniaria, Gorgonacea, etc., and increased food supply in others, e.g. Tridacnidae, Xeniidae, Foraminifera and possibly Radiolaria.

VI. SUMMARY

Intracellular association with unicellular algae is almost entirely confined to invertebrates which digest intracellularly. The mode of establishment must have varied according to whether the animal was herbivorous (or omnivorous) or a specialized carnivore. In the former a specialization of the algae to resist digestion would be the essential preliminary; stages in the establishment of such an association are indicated by conditions in the Spongillidae, *Frontonia*, *Paramecium bursaria* and Foraminifera (and possibly also Radiolaria). The more numerous cases of association with carnivores demand prior specialization of the animal to ingest and tolerate the algae; there is evidence that the initial stages were those of parasitism by the algae, indicated by conditions in various Hydrida. In *Chlorohydra viridissima* and the turbellarian *Dalyellia viridis* a balanced association has been attained. Further developments are represented by the animal becoming dependent on the plant (*Convoluta roscoffensis*), the plant on the animal (zooxanthellae of Coelenterata, etc.) and interdependence of the two (Xeniidae and other Alcyonacea). Infection by algae already specialized for life within other animals may explain the origin of the association in certain Gastropoda (e.g. *Aeolidiella alderi*), the Tridacnidae and some compound ascidians. Association may modify both partners in structure, physiology and habits. These may affect the individual or the race. The zooxanthellae are the most modified of the algae; the structure and habits of *Convoluta roscoffensis*, *C. paradoxa* and the Tridacnidae are most strikingly affected amongst the animals.

The significance of the association varies widely. Only rarely is it balanced, i.e. a true symbiosis. The algae gain protection and also inorganic food salts, notably those containing available nitrogen and phosphorus. The animals acquire an additional source of oxygen, but the value of this is uncertain. In some cases food in the form of organic matter is passed from the algae (e.g. *Paramecium bursaria*, *Convoluta paradoxa* and (in early life) *C. roscoffensis*) or the animal digests dead or living algae (instances of the former are somewhat uncertain but of the latter include the Spongillidae, *C. roscoffensis* in later life, Xeniidae and Tridacnidae). The great majority of coelenterates never digest associated zooxanthellae. The waste products of animal metabolism are always automatically removed by the algae. It is possible that the consequent gain in metabolic efficiency may be a significant factor in the great powers of growth and repair possessed by reef-building corals. Cases of experimental symbiosis illustrate the mutual advantage of the interrelationship. But the association ceases to be truly symbiotic when one party exploits the other, e.g. in *C. roscoffensis* (with eventual death of the animal) and, most efficiently of all because the animal continues to feed holozoically, in the Tridacnidae.

Algae are so frequently associated with tropical invertebrates as to constitute 'imprisoned phytoplankton' which, by tapping important sources of nitrogen and phosphorus, probably materially reduce the free phytoplankton of surrounding waters. The frequency of association in tropical waters may be attributed to the great competition there for the limited supplies of nutrient salts on the part of the plants, and to the greater survival value owing to increased nutriment (Tridacnidae, Xeniidae, Foraminifera, Radiolaria) or increased metabolic efficiency (Madreporaria and other Coelenterata) conferred on the animal.

VII. REFERENCES

Abe, N. (1939): *Palao Trop. Biol. Sta. Studies*, **1**, 651.
Ashworth, J. H. (1899): *Quart. J. Micr. Sci.* **42**, 245.
Berkeley, C. (1930*a*): *Quart. J. Micr. Sci.* **73**, 465. — (1930*b*): *Contr. Can. Biol. Fish.* **6**, 15.
Beutler, R. (1924): *Z. vergl. Physiol.* **1**, 1.
Bidder, G. P. (1920): *J. Linn. Soc. Lond. (Zool.)*, **34**, 305.
Bigelow, R. P. (1900): *Mem. Boston Soc. Nat. Hist.* **5**, 191.
Boschma, H. (1924): *Proc. K. Akad. Wet. Amst.* **27**, 13. — (1925*a*): *Proc. Nat. Acad. Sci. Wash.*, **11**, 65. — (1925*b*): *Proc. Amer. Acad. Arts Sci.* **60**, 451. — (1925*c*): *Biol. Bull. Woods Hole*, **49**, 407. — (1926): *Proc. K. Akad. Wet. Amst.* **29**, 993. — (1929): *Proc. 10th Int. Congr. Zool.* 920.
Boycott, A. E. (1926): *North-Western Nat.* **1**, 10.

BUCHNER, P. (1930): *Tier und Pflanze in Symbiose.* Berlin.
BUCHSBAUM, R. (1937): *Physiol. Zoöl.* **10**, 373.
CONKLIN, E. G. (1908): *Pap. Tortugas Lab.* **2**, 173.
DOYLE, W. L. & DOYLE, M. M. (1940): *Pap. Tortugas Lab.* **32**, 127.
ENDERS, E. E. (1909): *J. Morph.* **20**, 479.
FRITSCH, F. E. (1935): *The Structure and Reproduction of the Algae*, **1**. Cambridge.
FULTON, J. F. (1921): *J. Exp. Zool.* **33**, 353.
GAMBLE, F. W. & KEEBLE, F. (1904): *Quart. J. Micr. Sci.* **47**, 363.
GOETSCH, W. (1924): *Z. Morph. Ökol. Tiere*, **1**, 660.
GOETSCH, W. & SCHEURING, L. (1926): *Z. Morph. Ökol. Tiere*, **7**, 221.
GOHAR, H. A. F. (1940): *Publ. Mar. Biol. Sta. Ghardaqa*, **2**, 25.
GRAHAM, A. (1938): *Trans. Roy. Soc. Edinb.* **59**, 267.
HAECKER, V. (1908): Tiefsee-Radiolarien. *Wiss. Ergebn. Dtsch. Tiefsee-Exped. 'Valdivia'*, **14**.
HAFFNER, K. v. (1925): *Z. wiss. Zool.* **126**, 1.
HARDY, A. C. & GUNTHER, E. R. (1935): *Discovery Reports*, **11**, 1.
HARVEY, H. W. (1942): *Biol. Rev.* **17**, 221.
HOOD, C. L. (1927): *Biol. Bull. Woods Hole*, **52**, 79.
KAWAGUTI, S. (1937): *Palao Trop. Biol. Sta. Studies*, **1**, 107. — (1941): **2**, 319.
KEEBLE, F. (1908): *Quart. J. Micr. Sci.* **52**, 431. — (1910): *Plant Animals. A Study in Symbiosis.* Cambridge.
KEEBLE, F. & GAMBLE, F. W. (1907): *Quart. J. Micr. Sci.* **51**, 167.
LIPMAN, C. B. & TEAKLE, L. J. H. (1925): *J. Gen. Physiol.* **7**, 509.
LOEFER, J. B. (1936*a*): *J. Exp. Zool.* **72**, 387. — (1936*b*): *Amer. Nat.* **70**, 184. — (1938): *Arch. Protistenk.* **90**, 185.
MARSHALL, S. M. (1932): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 253.
MAYOR, A. G. (1918): *Pap. Tortugas Lab.* **12**, 175. — (1924): **19**, 1.
MOUCHET, S. (1930): *Sta. Ocean. Salammbô*, Notes, no. 15.
NAVILLE, A. (1926): *Rev. Suisse Zool.* **33**, 251.
PARKER, R. C. (1926): *J. Exp. Zool.* **46**, 1.
PASCHER, A. (1930): *Biol. Zbl.* **50**, 1.
PRATT, E. M. (1906): *Quart. J. Micr. Sci.* **49**, 327.
PRINGSHEIM, E. G. (1928): *Arch. Protistenk.* **64**, 289.
PÜTTER, A. (1911): *Z. allg. Physiol.* **12**, 297.
RUSSELL, E. S. (1927): *Biol. Rev.* **2**, 213.
SMITH, H. G. (1935): *Ann. Mag. Nat. Hist.* (10), **15**, 615. — (1936): *Pap. Tortugas Lab.* **31**, 17. — (1939): *J. Exp. Biol.* **16**, 334.
THIEL, M. E. (1929): *Zool. Anz.* **81**, 295.
TRIGT, H. VAN (1919): *Tijdschr. Ned. Dierk. Ver.* (2), **17**, 1.
UCHIDA, T. (1926): *J. Fac. Sci. Imp. Univ. Tokyo* (IV), **1**, 45.
VERWEY, J. (1931): *Treubia*, **13**, 169.
WELSH, M. (1936): *Biol. Bull. Woods Hole*, **70**, 282.
WILSON, D. P. (1932): *Philos. Trans.* (B), **221**, 231.
YONGE, C. M. (1931*a*): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 83. — (1931*b*): *Nature, Lond.*, **128**, 309. — (1934): **134**, 12. — (1935): *Proc. Linn. Soc. Lond.* Sess. 147, 90. — (1936): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 283. — (1937*a*): *Biol. Rev.* **12**, 87. — (1937*b*): *Pap. Tortugas Lab.* **31**, 207. — (1940): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 353.
YONGE, C. M. & NICHOLAS, H. M. (1940): *Pap. Tortugas Lab.* **32**, 287.
YONGE, C. M. & NICHOLLS, A. G. (1931*a*): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 135. — (1931*b*): **1**, 177.
YONGE, C. M., YONGE, M. J. & NICHOLLS, A. G. (1932): *Sci. Rep. Gt Barrier Reef Exp. Brit. Mus.* (*Nat. Hist.*), **1**, 213.

THE STRUCTURE AND PERMEABILITY OF BLOOD CAPILLARIES

By J. F. DANIELLI and A. STOCK
From the Departments of Zoology and Biochemistry, Cambridge University, and the Department of Zoology, Bedford College for Women, London University

(*Received* 27 *March* 1944)

In this review we shall deal mainly with two problems which are functionally closely related, namely, the structure and the permeability of the blood capillary wall. In connexion with structure we shall pay attention to the cells responsible for opening and closing the capillaries. The school of Krogh, on the one hand, has emphasized the importance of the Rouget cells. On the other hand, American experimental workers, and in particular German histologists and pathologists, have laid emphasis on the ability of the endothelial cells themselves to carry out the vital role of modifying the lumen of the vessels. In the field of permeability we shall deal particularly with Starling's hypothesis of ultrafiltration and resorption of fluid through the capillary wall. This hypothesis has been accorded general recognition and a considerable measure of proof has been provided by the work of Landis. Rous and his colleagues, however, have brought forward a substantial body of evidence which appears to be incompatible with Starling's hypothesis. Further evidence is now available, in both the anatomical and physiological fields, which goes a great way, at least, towards resolving these differences of opinion.

I. THE CONTRACTILE CELLS

Histology of the capillary wall. From the point of view of their structure blood capillaries have been defined as vessels consisting of a single layer of endothelial cells. In the capillaries of most organs the boundaries of these cells can be demonstrated by impregnation with silver salts. In some few organs, and in growing capillaries, the impregnation fails to produce the silver markings (Clark & Clark, 1939), and this has been regarded by some authors as evidence for the existence of an endothelial syncytium. Observations by Schaffer (1920), Volterra (1925*a*, *b*), Huzella (1925), Plenk (1927, 1930), Pfuhl (1933), and Loeschke & Loeschke (1934) have shown that an outer tube surrounds the endothelium of the capillary. This outer tube consists of a network of argyrophil fibres originating from branches of the connective tissue in the organ through which the capillary runs (Fig. 1). The meshes of this lattice-work, in the opinion of some workers, are filled with a delicate membrane, providing a continuous sheath around the endothelium. Between the endothelium and this outer tube ('Grundhäutchen' of Plenk, 'reticulo-adventizia' of Volterra, perithelium) is situated a lymph space (Pfuhl, 1933; Heimberger, 1927*a*), and it is thought that there are fibrous connexions ('Basalfransen', Zimmermann, 1923) bridging this space. Embedded in or attached to the surface of this connective tissue tube there appear to be situated branched cells, the branches of which surround the capillary (Fig. 2). Plenk, Volterra, Pfuhl, and Loeschke & Loeschke, on the

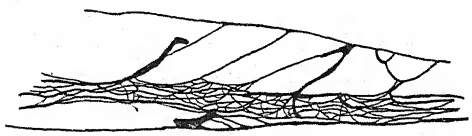


Fig. 1. Argyrophil fibres around a capillary in the psoas muscle of the guinea-pig. Redrawn from Nagel (1934).

one hand, regard the perithelium as an essential part of the capillary; they take the view that the branched cells are identical with cells variously termed 'adventitia cells' (Marchand, 1923), 'pericytes' (Zimmermann, 1923), or 'Rouget cells' (Vimtrup, 1922). Some observers have regarded these cells as being contractile and responsible for the contractions of the capillary. Krogh & Vimtrup (1932),

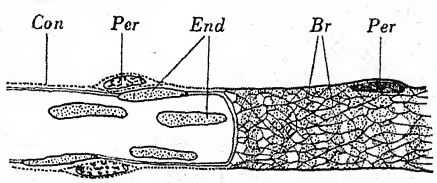


Fig. 2. Diagram of a capillary from the mesentery of a vertebrate animal, redrawn from Pfuhl (1934). *Br*, branches of pericyte; *Con*, connective tissue tube of the capillary; *End*, endothelial nucleus; *Per*, nucleus of a pericyte.

however, are of the opinion that the perithelium does not belong to the capillary but to the surrounding organ and that the Rouget cells are directly attached to the endothelium.

Contraction of the capillaries. Observers of the contraction of capillaries have described two essentially different modes by which the narrowing of the lumen can be effected. The first mode, observed by Stricker (1865, 1879), Golubew (1869),

Tarchanoff (1874), and confirmed by Steinach & Kahn (1903), consists of a swelling of endothelial nuclei which thereby narrow or even obliterate the lumen. Recently, Kahn & Pollak (1931) have confirmed these earlier observations. They have produced microphotographs of capillaries of the nictitating membrane of frog, in which the lumen appears to be obstructed by endothelial nuclei bulging into it; they point out that the external diameter of the vessels remains unaltered. Similar results were also obtained by Field (1935), and indirectly by Heimberger (1927*a*), who describes localized one-sided obstructions of the lumen of capillaries in the nail bed of man occurring after mechanical stimulation of the capillary wall. He attributes the obstructions to swelling of endothelial nuclei. Owing to the thinness of the capillary wall no corresponding changes in the endothelial cytoplasm have been conclusively demonstrated (Barksdale, 1926),

The unaltered external diameter of capillaries whose lumen has been reduced by 'nuclear swelling' has led some observers to assume that imbibition processes were responsible for the alteration in the width of the lumen. Kahn & Pollak (1931) point out that the passive swelling of the nuclei must be due to a process fundamentally different from, for example, the contraction of a smooth muscle fibre. But another and perhaps more likely explanation of the nature of the narrowing of the lumen has been advanced by other authors. Heimberger (1927*a*), for instance, comes to the conclusion that a contraction of the endothelial cells as a whole produces the intrusion of endothelial nuclei into the lumen, and he argues that imbibition processes would fail to produce such rapid contractions as he observed. There are many other indications showing that the endothelium possesses contractility. Observations of contractions in the absence of, or in places remote from, possible extracapillary contractile elements point this way. Thus Clark & Clark (1925*b*) describe contractions in the ingrowing capillaries of the tadpole's tail independent of, or in advance of, extracapillary contractile cells; these are apparently not only due to 'nuclear swelling' but involve also an 'active contraction', by which the capillary assumes a wavy outline, the external diameter being thus altered too. Florey & Carleton (1927) arrived at similar results on histological grounds. Sandison (1932) also concludes that the endothelium possesses a slight contractility.

The fact that the capillary endothelium has been shown to be contractile has led Jones (1936) to the view that elements described by histologists as endothelial cells with outlines demonstrable by silver salt solutions are, in fact, smooth muscle cells. Derived from the circular muscle coat of the arterioles, they become spindle-shaped, assume a spiral form and overlap each other at the edges. The silver lines are then due to a reticulum, covering the muscle fibres and holding them together. There is therefore 'no need to invoke any exceptional apparatus in order to explain capillary constriction'. Michels (1936) finds nothing to confirm Jones's assumption. In his preparations there is no evidence for arteriolar circular muscles becoming fusiform and spirally twined endothelial cells. He, too, advances the view that in the web of the frog's foot 'under normal and experimental conditions capillary caliber changes may be effected by the regional endothelium...'.

The second mode of contraction has been described as a narrowing of the lumen with corresponding changes in the external diameter of the capillary, whereby the wall of the vessel becomes longitudinally folded. The existence of such folds has been strongly maintained by Steinach & Kahn (1903), Vimtrup (1922, 1923), Field (1935) and others, and more recently by Kahn & Pollak (1931), who demonstrate them on microphotographs of capillaries of the frog. They have been inferred by Heimberger (1927*a*), who observed that in capillaries of the human nail fold, stimulation by electric current at a single point of the wall caused a splitting of the blood thread into narrow channels; this, in his view, strongly indicates a folding of the capillary wall. Such a folding could only be explained in one of two ways: either by the collapse of the capillary due to reduced pressure in the arterioles, or by a constriction due to extra-capillary elements. The first explanation is very difficult to maintain. Not only is it improbable that such a reduction in pressure can have an effect in a very limited stretch of one single capillary, but also the capillary wall possesses, as Nagel (1934) has demonstrated, extraordinary elasticity when pulled by a micro-dissection needle. Reduced pressure could, perhaps, reduce the capillary lumen by a thickening of the elastic wall, but cannot explain its folding. Krogh (1936) has also advanced arguments against the view that capillaries are passively distended by the pressure of the blood. The only other explanation so far suggested of this folding of the wall, observed and confirmed by a number of authors, lies in the activities of extra-capillary contractile elements.

Rouget (1873, 1879) observed in the nictitating membrane of the frog branched cells on the capillary wall, which he regarded as contractile. Later Mayer (1899, 1902) again described such cells, which were the object of detailed studies by Zimmermann (1923) and Vimtrup (1922, 1923). The latter observed their contraction, which constricted the capillary lumen, and termed them 'Rouget cells'. It was assumed that the smooth muscle cells undergo a gradual transformation from the typical fusiform shape, as found on the arterioles, into branched cells, the nuclei of which arrange themselves longitudinally on the capillary, and the branches of which tend to encircle it (Fig. 3). In action the branches contract and the formerly flat nuclei become pro-

minent, protruding towards the outside as well as towards the capillary lumen.

Vimtrup's observations had led to a controversy not only on the manner of the capillary contraction but also on the nature of these peri-capillary cells. The difficulty is that the nucleus alone can be stained by ordinary histological methods, the cytoplasm remaining unstained. Only methylene blue has so far resulted in a satisfactory staining of the whole cell. Moreover, the perithelium of the capillary does not consist of just one type of cell. Fibrocytes, histiocytes and nuclei of Remak fibres (Schwann's cells) become attached to the connective tissue network around the endothelium. They all assume a flattened shape and therefore in sections are easily mistaken for Rouget cells. So far, characteristic differences between nuclei of Rouget cells and those of other perithelial cells have not conclusively been demonstrated. The results of the work carried out by histologists and pathologists, such as Marchand (1923), Benninghoff (1926, 1930), Urtubey (1932), Pfuhl (1933), Loeschke & Loeschke (1934) and others, has to be considered in the light of those

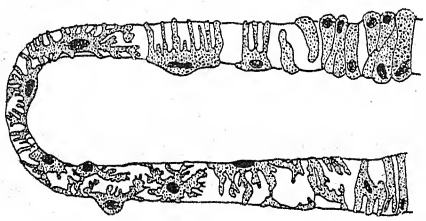


Fig. 3. Transition of vertebrate smooth muscle fibres into capillary pericytes. Redrawn from Benninghoff (1930).

difficulties. Clark & Clark (1925*a*), Clark, Clark & Abell (1933) and others have shown that, in the development of capillaries, connective tissue cells become attached to the capillary wall, and assume a flattened shape, becoming similar to Rouget cells. Vimtrup (1923). Bensley & Vimtrup (1928) and Defrise (1929) have been able to demonstrate fibrils, which they regard as myofibrils, in the Rouget cells. Busch (1929) has claimed to have obtained evidence of their innervation, but Michels (1936) and Jones (1936) concluded that in the cases investigated by them the cells involved were those of Remak fibres; this makes a reinvestigation of Blum's work desirable.

Complications involved in the views of Vimtrup, Krogh, and Bensley & Vimtrup have been brought out by Benninghoff (1926) and by Nagel (1934). Benninghoff points out that the contraction of the protoplasm around the nucleus of the Rouget cells, which results in the nuclear extrusion described above, would also produce tension in a longitudinal direction, since the cell body is arranged longitudinally on the capillary wall; this longitudinal tension would tend to counteract the radial tension set up by the contracting branches. Zimmermann (1923) and also Vimtrup (1923) sought to resolve this difficulty by assuming that only the branches were capable of contraction. But this assumption does not explain the changes which can in fact be observed in the whole of the contracting Rouget cell. Nagel's investigations demonstrated an extreme elasticity of the endothelium which suggests that contractions of Rouget cells not completely surrounding the capillary would tend to stretch the capillary wall, and would be but partly effective in narrowing the lumen.

The difficulties prescribed by these observations might be resolved by the assumption that the Rouget cells are not directly attached to the endothelium, but to the connective tissue fibres around it. The picture would then essentially resemble the position of smooth muscles embedded in their own reticular sheath. Very little is known about the mechanical properties of these argyrophil fibres. Plenk (1927) regards them as rigid structures, while Huzella (1925) ascribes to them a rubber-like elasticity. This latter view is denied by Plenk in later papers (1930, 1937). Volterra (1925 *a*, *b*, 1933), Fieschi & Storti (1929) and Volterra & Schupfer (1934) are inclined to regard the reticular fibres of the capillary as solely responsible for true contraction. They envisage a mechanism whereby, owing to changes in hydrogen-ion concentration, swelling and contraction of these fibres produce constriction of the capillary.

We therefore come to the conclusion that, unless one of these groups of workers is totally mistaken, the constriction of capillaries can be achieved in several different ways. A few endothelial cells only may be involved, reducing the capillary lumen by intrusion of nuclei without change in the external capillary diameter; or activity of a greater stretch of endothelial cells may result in constriction of the capillary without folding of the wall. But besides that, constrictions with a folding of the wall have been demonstrated. They can best be explained by the action of extra-capillary elements embedded in the connective tissue sheath of the capillary. It is probable that these cells are the Rouget cells, though a clearer histological definition to distinguish them from connective tissue and nervous elements, also included in the perithelium, is still greatly to be desired.

II. THE FACTORS AFFECTING THE RATE OF FILTRATION THROUGH THE CAPILLARY WALLS

Before we can proceed to discuss the Starling hypothesis we must consider the many factors which affect capillary permeability, since these will, of course, have to receive due attention in the analysis of experimental results. In so doing we shall assume that Starling's hypothesis is correct, so as to avoid breaking up the text with innumerable reservations.

First, we must define capillary permeability. Landis (1927) defined capillary permeability as the volume of fluid filtered across unit area of capillary membrane per unit pressure difference across the membrane per unit time, e.g. μ^3/μ^2/atmosphere/min. Here the pressure difference is that between the blood pressure in the capillary, on the one hand, and the net colloid osmotic pressure difference across the capillary membrane on the other hand. This is the ideal unit. But in many practical cases, where it is desired to deal with a relatively normal tissue, the capillaries are not open to microscopic observation, so that changes in capillary diameter cannot be measured. Consequently in such studies changes in capillary membrane area cannot be measured, and it is necessary to define capillary permeability as the volume of fluid filtered across the capillary membrane per unit length of capillary per unit pressure difference per unit time (Danielli, 1940). The subsequent discussion will be in terms of this second unit unless otherwise specified.

1. *Mechanical factors*

The hydrostatic pressure on the wall of the capillary. Variations in this may occur as the result of changes in the action of the heart (or in the head of pressure used for perfusion), as the result of constriction or dilatation of the other vessels, including capillaries and capillary loops, and by the opening of more capillaries leading from the same arteriole to the same venule, or connected by anastamoses to other arterioles and venules. In the normal animal, judging from the results of Landis (1934), and from rates of lymph formation (Drinker & Joffey, 1942), this hydrostatic pressure slightly exceeds the colloid osmotic pressure of the blood, leading to a slow net formation of lymph which is returned to the blood stream by the lymphatic circulation. When the lymphatics become blocked, or the blood pressure rises to excessive heights so that the flow of lymph is too great to be dealt with by the lymphatic system, oedema develops.

Colloid osmotic pressure. Other things being equal, increase in the colloid osmotic pressure of the blood will lead to less filtration across the capillary wall. Clinically a deficiency in the colloid osmotic pressure, such as is found with marked albuminurea, dietary deficiency in protein and other conditions, leads to generalized oedema. The colloids of the blood are only effective in opposing the hydrostatic pressure in the capillaries in so far as the size of the pores in the capillary walls is such as to prevent passage of the colloid into the peri-capillary spaces. Consequently a colloidal molecule of small diameter, such as that of inulin, is not very effective in maintaining the colloid osmotic pressure of the blood, whereas a comparatively large molecule, such as that of serum globulin, is very effective. Krogh & Harrop (1921) found that 3 % gum acacia, which has a colloid osmotic pressure, when measured in a collodion sac, roughly equal to that of frog plasma, does not prevent the development of oedema in perfusion experiments with frogs. This observation was confirmed by Drinker (1927). Krogh & Harrop found that the addition of defibrinated ox blood to acacia prevents oedema formation and Drinker found that 20 % horse serum added to acacia prevented oedema formation. Saslow (1938), using a different species of frog, could not confirm Drinker's results. Danielli (1940) found that highly purified ovalbumin and haemoglobin are both relatively ineffective in preventing oedema formation, that acacia is rather more efficient and that sheep, horse and ox serum are all very efficient. Addition of 10 % of ox serum to gum acacia (with appropriate addition of water to bring the resulting mixture to roughly the same colloid osmotic pressure and ionic strength as frog plasma) is sufficient to reduce the rate of development of oedema to 4 % of that found when gum acacia is the only colloid present. The serum used in these experiments was incubated at 50° C. for a short time to destroy vasotonins and then filtered through a Seitz filter. Such serum is free of all particulate matter and contains little, if any, vasoconstrictor substances. When used as the only colloid in a solution of the same colloid osmotic pressure as frog plasma, such serum is still able to reduce the rate of oedema formation in frogs to a value of the order of 1 % of that found when gum acacia is the sole colloid present.

The number of capillaries open. It is obvious that if no capillaries at all are open there can be no filtration through the capillary wall. If, on the other hand, a large number of capillaries are open, the hydrostatic pressure in the capillaries may fall below the colloid osmotic pressure, and so again no fluid will be filtered across the wall, but may, on the contrary, be withdrawn from the tissue spaces into the circulation. Thus, on passing from a state where few, or no, capillaries are open to a state where all the capillaries in a given tissue are open, the total rate of filtration across the capillary walls may rise to a maximum and then fall again, possibly reaching a negative value.

The area of capillary wall available for filtration. Considering extreme cases, a given flow of fluid through the capillaries may be achieved in either of two ways: a considerable number of capillaries may be open, but have a small diameter, or a small number of capillaries may be open, but be dilated. Since the rate of flow through a cylindrical tube varies as the fourth power of its radius, whereas the area of its surface varies only as the first power of its radius, it is clear that in the first case mentioned the area available for filtration will be much greater than in the second case.

The intercellular cement and the pore size. All the evidence available goes to show that the capillary wall in a normal tissue is comparatively impermeable both to serum albumin and to serum

globulin. Consequently, the average effective pore size must be rather less than 6 mμ. It therefore follows that the pores cannot simply be the spaces between adjacent endothelial cells, as these spaces are very much greater than 6 mμ. The pore size is probably defined by the intercellular cement lying between the endothelial cells. The importance of this cement was emphasized by Overton and the cement has been studied in other connexions by Herbst, by Gray and by E. B. Harvey. It is known that the cement dissolves in the absence of calcium or at slightly acid *p*H, and recently Chambers & Zweifach (1940) have shown that when frog tissues are perfused with calcium-free solution the intercellular cement dissolves and capillary permeability is enormously increased, so that, whereas in the presence of sufficient calcium and at neutral *p*H the capillary wall is almost impermeable to protein and quite impermeable to carbon particles, when the cement is dissolved the carbon particles can be seen to be forced in jets of fluid through the spaces between adjacent endothelial cells. It is quite certain, therefore, that anything which affects the physical condition of the intercellular cement will affect the pore size and this in turn will affect the amount of filtration across the capillary wall. The exact physical nature of this intercellular cement is unknown, but it seems very likely that under normal conditions the walls of the pores through this cement are coated with a layer of adsorbed serum protein which effectively reduces the area of pore through which filtration may occur. Displacement of this adsorbed serum protein by other molecules of small diameter may result in an increase in pore size. This may be one reason why, for example, the glomerular capillaries, which are normally impermeable to serum albumin, become permeable both to serum albumin and to ovalbumin when ovalbumin enters the blood stream.

It is possible that part of the increase in capillary permeability sometimes found in the reaction to foreign proteins is due to a combination of the antigen with antibody adsorbed on the walls of the pores. Duran-Reynals (1939) claims that the spreading factor of McClean (1930) causes an increase in capillary permeability which is presumably due to an action upon the intercellular cement. In this case the intercellular cement is presumably a mucoprotein. Danielli (unpublished) has tried to check this observation in the perfused frog, using mammalian spreading factor, but could find no increase in permeability such as could be due to enzyme action on the intercellular cement, although there is some increase in rate of filtration through the capillary wall due to arteriolar vaso-dilation, the preparations containing a vaso-dilator substance. However, this point cannot be settled until homologous spreading factor is available, as it is quite possible that mammalian spreading factor will not act upon frog mucoprotein.

2. *Plasma and hormone factors*

Particulate matter. Since we are dealing with filtration through a pore system it is clear that particles of an appropriate shape and size will be able to block the pores and so reduce the permeability of the capillary wall. Krogh (1929) found that addition of ox red cells to gum acacia greatly restricted the development of oedema in perfused frogs without at all changing the calibre of the vessels. This was confirmed by Saslow (1938) who attributed the action of the red cells to the enhanced supply of oxygen they make available to the tissues. Since, however, both Drinker (1927) and Danielli (1940) found that an impermeable state of the capillaries resulted when serum alone is added to the acacia, the action of the red cells cannot be due to their oxygen-carrying capacity. On the other hand, addition of small amounts of blood platelets to gum acacia reduces the permeability almost as effectively and, weight for weight, far more effectively, than does filtered serum. As these platelet preparations have no vaso-motor action it seems necessary to attribute their effect to mechanical blocking of the pores. Red cell ghosts were found to have a similar effect, though neither red cells nor their ghosts were as effective as blood platelets. Chambers & Zweifach (1940) found that red cells and suspended carbon particles both reduced the rate of oedema in perfused preparations and also attributed their results to mechanical blocking of the pores.

In view of these observations on frogs the function of platelets in the control of capillary permeability in normal mammals is of some interest. Some types of purpura are known to be closely associated with platelet deficiency, and Bedson (1922) has shown that after administration of an anti-platelet serum the blood platelet count is reduced to a very low figure and the typical capillary fragility of purpura sets in. On the other hand, an oedematous condition is not observed. Feldberg & Danielli (unpublished) have studied the effect of replacing the blood of living cats by red cells added to Seitz-filtered (platelet-free) cat serum. They found no signs of increased permeability after removal of more than 80 % of the blood platelets. It therefore seems probable that the blood platelets constitute a reserve mechanism. Normally the blood capillaries are sufficiently impermeable to protein for the platelet mechanism not to come into play; moderate damage to a capillary, such as may occur in the course of daily life, produces an increased rate of filtration as a result of which platelets are sucked across the damaged region, and both restrict the loss of fluid through the capillary wall and prevent the loss of red cells through the damaged area. When blood platelets are reduced in amount red cells get sucked into abnormally permeable areas in place of platelets and, owing to their great ease of deformation, tend to pass through the capillary wall. Thence arises the

purpuric condition. In this condition the capillaries are maintained relatively impermeable even when damaged, at the expense of loss of red cells into the tissue spaces.

Hormones, etc. The hormones and hormone-like bodies known to affect capillaries are already numerous. They include histamine, adrenaline, pituitrin, Lewis's H substance, leucotaxin and the adrenal cortical hormone. Histamine and adrenaline were studied by Dale & Richards (1919) who found that histamine has a dilator action on the capillaries of the cat which can only be observed in the presence of sufficient adrenaline to give a tone to the vessels. On the other hand, Clark (1934) found that adrenaline causes dilation of muscle capillaries. Krogh (1924) found that pituitrin causes constriction of the capillaries. Lewis (1927) found extensive evidence for the release of a substance in the skin by injury which causes capillary dilation. This H substance is certainly not identical with histamine. Menkin (1940) has isolated a substance from inflammatory exudates, leucotaxin, which causes capillary dilatation and to which leucocytes are positively chemotactic. In these respects leucotaxin closely resembles Lewis's H substance. Swingle *et al.* (1933, 1934) have shown that the adrenal cortical hormone to some extent prevents the increase in capillary permeability found in shock. Hyman & Chambers (1943) found that the adrenal cortical hormones are extremely efficient in reducing capillary permeability, their action being pronounced at a dilution of $1:10^8$. They have no dilator or constrictor action, so far as could be observed, and so presumably affect the form of the endothelial cells and thus modify the available area (thickness or pore size of the intercellular cement). Menkin (1940) and Freed & Lindner (1941) found that the increase in capillary permeability caused by leucotaxin in rabbits is antagonized by the adrenal cortical hormone. It is likely that all these hormones play some part both in maintaining the normal impermeability of the capillary and in the development of shock-like conditions. Substances like leucotaxin and Lewis's H substance are particularly likely to cause complications in the perfusion of tissues, and the absence of such substances as the cortical hormones from artificial perfusates may also lead to misleading conclusions.

Metabolites, etc. Krogh (1924) and Landis (1927) found that oxygen lack caused capillary dilatation and increase in capillary permeability. Hemingway & McDowell (1927) found that in cats there is a loss of capillary tone in perfused limbs due to accumulation of lactate. No doubt many other metabolites also have a vaso-dilator action. Fleisch & Weger (1937) and Fleisch (1937) have studied numerous metabolic products, including adenosine triphosphate, which has a vaso-dilator action in moderately low concentrations, of the order of $1:2{\cdot}5 \times 10^{-5}$. These results are of the more interest in view of evidence, provided by Green (1943) and Bielschowsky & Green (1943), that adenosine triphosphate or a related compound is a primary agent in ischaemic (crushed limb) shock. Such factors are likely to be of importance both in shock and in perfused preparations. Raising the potassium concentration to ten times the normal level has no effect on capillary permeability in the perfused frog (Danielli, unpublished). Diminution in calcium content or of the *p*H increases capillary permeability by increasing the solubility of the intercellular cement (Chambers & Zweifach, 1940).

3. *Local tonus and nervous control*

Nervous control. The blood capillaries are known to receive both non-medullated sympathetic fibres and medullated fibres belonging, for example, to the same nerve network which innervates the sense organs of the skin. Steinach & Kahn (1903) found that stimulation of the sympathetic system causes constriction of the capillaries of the frog's nictitating membrane. Krogh, Harrop & Rehberg (1922) found that sympathetic activity caused constriction of frog skin capillaries. Similarly, sympathetic stimulation causes constriction of human skin capillaries (Leriche & Policard, 1920), of frog muscle capillaries (Gabbe, 1926), of rabbit ear capillaries (Krogh, 1920; Harris & Marvin, 1927; Beecher, 1936) and of cat's ear capillaries (Hooker, 1920). But Hartman, Evans, Melachowski & Michalek (1928) found that, in muscle, sympathetic stimulation causes dilatation of the muscle capillaries. As Clark (1934) found that adrenaline causes dilation of muscle capillaries there may be a fairly general unity of sympathetic activity: where sympathetic activity causes constriction of arterioles, as in the skin, it is likely also to cause capillary constriction: where arteriolar dilatation is found, as in the skeletal muscles, capillary dilatation is also to be expected.

Doi (1920) and Krogh *et al.* (1922) found that stimulation of posterior nerve roots causes dilation of many of the capillaries of the skin of the frog, but that there was little evidence of any change in the capillaries of frog muscle, which receives relatively few sensory (medullated) fibres. Lewis & Marvin (1927) similarly found stimulation of the posterior roots causes dilatation of the capillaries of the cat's foot. There is some doubt as to whether in the normal animal these so-called antidromic impulses can pass from the central nervous system along the sensory fibres, but reflex dilatation of the skin capillaries, due to stimulation of the skin sense organs, is likely to be of common occurrence and may again be of some importance in perfusion experiments.

A number of studies have been made of the effect of sympathectomy on capillary permeability. In such studies it is very difficult to separate effects due to the changes in calibre of the various vessels following sympathectomy from a true change in

permeability of the capillary wall. These results will be discussed later.

Intrinsic tonus. It is likely that both the endothelial cells and the Rouget cells have powers of maintaining a tonus in the absence of any nervous and possibly of any humoral control. For example, after sympathectomy there is an initial dilatation of both the arteries and capillaries of the rabbit's ear, but after about 24 hours the capillaries begin to constrict again, although little change is observable in the major vessels. When a frog is perfused with gum acacia solution containing no vaso-motor substances all the vessels, including the capillaries, are initially widely dilated, but after a short time of perfusion some degree of tone is recovered and in perfusion experiments it is necessary to wait until this recovery has occurred before a steady state is reached.

4. *Dietary factors*

The investigation of the influence of dietary factors on capillary permeability is still in its infancy. Deficiency of vitamin C, vitamin K or vitamin P may lead to increased fragility of the capillaries, and one form or another of purpura results. It is not yet known whether the essential change in these conditions is in the endothelium of the capillaries, in the effectiveness of the platelets in repairing local damage, or in some other plasma factor. But in all these conditions an increase in permeability is involved. A well-known cause of increased filtration through the capillaries is dietary deficiency of protein, in which condition the protein content of the blood and therefore its colloid osmotic pressure are reduced, so that generalized oedema appears. Although the importance of really adequate diet for experimental animals has not been given much attention in the past in connexion with capillary studies, it is likely that attention to this point would be profitable.

III. STARLING'S ULTRAFILTRATION HYPOTHESIS

According to Starling (1895) the rate at which fluid is filtered off from the blood plasma through the capillary wall is partly determined by the difference between the hydrostatic pressure in the capillary and the colloid osmotic pressure of the plasma. Over the arteriolar end of a capillary, it was argued, the hydrostatic pressure was likely to exceed the colloid osmotic pressure, and consequently fluid would pass from the blood stream into the peri-capillary spaces. Over the venular end of the capillary it was thought likely that the hydrostatic pressure would be less than the colloid osmotic pressure and consequently fluid would be recovered from the peri-capillary spaces into the blood stream. There must consequently be a flow of peri-capillary fluid from the arteriolar to the venular end of a capillary. Landis (1934) has summarized the evidence showing that the blood pressure in the capillaries does indeed follow the requirements of Starling's hypothesis. If the capillary wall is completely impermeable to protein, the whole of the fluid filtered off over the arteriolar region may be recovered over the venular region. This is not the case if the capillary has even a slight permeability to protein. Let us suppose that the filtrate from the blood stream contains roughly 1 % of the protein content of the plasma. As this filtrate moves in the peri-capillary space down to the venular end of the capillary, fluid will be withdrawn from it into the plasma, but just as in filtration from the blood stream only 1 % of the initial protein content of this fluid passed with it into the peri-capillary spaces, so over the region of resorption only 1 % of the protein of the peri-capillary fluid will be returned to the blood stream. The remaining protein will be concentrated, increasingly so on passing to the extreme venular end of the capillary, where its concentration may even approach that of the protein in the blood plasma. There must thus be a gradient in the concentration of colloidally active material present in the peri-capillary fluid in passing from the venular to the arteriolar end of the capillary. The fluid of highest protein content, found at the venular end, cannot be recovered into the blood stream directly, but is returned by the lymphatic circulation.

Criticisms of this general mechanism of filtration and recovery of fluid into the capillaries have been made, based on evidence adduced by Drinker & Field in one case, and by Rous and his colleagues in the other.

Drinker & Field (1933), in their book on the lymphatic circulation, show that the lymph fluid has a protein content of the order of 1 % or more, and argue that this protein content in the peri-capillary fluid must increase the amount of filtration by reducing the net colloid osmotic pressure difference across the blood capillary membrane, which is the force retaining fluid in the blood stream. If this argument is taken at its face value, it would mean that with capillary blood pressures of the order of magnitude of those reported by Landis (1934) there would be little or no recovery of fluid over the venular end of the capillaries. This argument, however, is not correct. It fails to appreciate that a protein content of 1 % in the lymphatic fluid corresponds to a protein content of 1 % at the extreme venular end of a capillary and consequently, as explained above, to a much lower protein content at the arteriolar end of the capillary.

The second set of arguments against the Starling hypothesis is based on observations made with dyes by Rous, Gilding & Smith (1930), Smith & Rous (1931 *a*, *b*), Rous & Smith (1931), McMaster, Hudack & Rous (1932), and Smith & Dick (1932). The facts are as follows. When a readily diffusible dye is introduced into the blood stream it appears in the peri-capillary fluid in roughly the same

concentration as in the blood stream, and there is no gradient of concentration in the peri-capillary fluid on passing from the arteriolar to the venular end of the capillary. With poorly diffusible dyes only a small proportion of the concentration in the blood stream appears outside the capillary; the concentration may be so low as to be negligible in the arteriolar peri-capillary fluid, rising to a relatively high value outside the extreme venular capillaries and small veins. This distribution of poorly diffusible dye is taken by Rous and his colleagues to indicate that there is a gradient of permeability on passing from one end of the capillary to the other, the permeability reaching a maximum value at the extreme venular end and in the small veins. If this contention were correct the process of recovery of fluid over the venular end of the capillary would be gravely hampered, if not entirely abolished, by the loss of colloid from the blood stream in this region. But this is not the only possible interpretation of the facts. The dyes which are classed as highly diffusible are those which exist as molecules of relatively low molecular weight and which are in true solution; those which are classed as poorly diffusible form colloidal solutions and also are rather strongly adsorbed on the plasma colloids. These poorly diffusible dyes must thus be regarded as one of the colloidal constituents of the blood: as they are strongly adsorbed on the plasma protein and often only a minute fraction of the total plasma concentration of dye is present as single molecules in solution, it follows that where a high concentration of such a dye appears outside the blood capillaries there must be a high concentration of colloid, and in particular there is likely to be a high concentration of protein. The appearance of a gradient of concentration of poorly diffusible dye increasing from the arteriolar to the venular end of the capillary must therefore involve a concentration gradient of colloidal molecules, and almost certainly of protein molecules, over this region. Now in the first part of this section it was pointed out that given a small degree of permeability of the blood capillaries to protein (and this is an experimental fact) it follows from Starling's hypothesis that there must be a gradient of protein content in the peri-capillary fluid, such that the protein concentration reaches a maximum outside the extreme venular end of the capillaries and small venules. Consequently when a dye is present in the blood stream which is strongly adsorbed upon the plasma proteins, so that most of the dye is in the adsorbed condition, it follows from Starling's hypothesis that there will be a gradient of dye adsorbed upon protein in the peri-capillary fluid of exactly that character which is found by Rous and his colleagues. In fact, the observations of Rous, far from constituting an objection to Starling's hypothesis, are a necessary consequence of the mechanism proposed by Starling, and may be regarded as additional evidence that qualitatively the hypothesis is correct.

Most of the studies designed to test the accuracy of Starling's hypothesis have involved measurement of the rate of development of oedema or of filtration across the capillary wall. Landis (1927) introduced a technique whereby the rate of filtration across the wall of a single blood capillary could be measured. The steps involved in this technique are the following: (*a*) Blockage of the venous end of a capillary by pressure from a micro-dissection needle. Attention is then focused on an erythrocyte proximal to the blockage; as fluid is filtered off through the blood capillary wall under the hydrostatic pressure in the capillary the red cell moves slowly down the lumen. The rate of loss of fluid can be calculated by measuring the rate of movement of the erythrocyte, the length of capillary between the erythrocyte and the block, and the radius of the capillary. (*b*) Measurement of the hydrostatic pressure in the capillary. This is done by canulating the capillary with a micro-canula and observing the hydrostatic pressure which is just sufficient to prevent blood entering the canula. From measurements of this type Landis concluded that the rate of filtration varied linearly with the hydrostatic pressure in the capillary. When the hydrostatic pressure was greater than the effective colloid osmotic pressure of the blood, fluid was lost to the peri-capillary spaces; when the hydrostatic pressure was less than the effective colloid osmotic pressure, fluid was regained from the peri-capillary spaces. Wind (1937) has also made studies by Landis's technique. The chief criticisms that have been made of the results obtained by Landis are: that the tissues are subjected to considerable manipulation and are in a somewhat abnormal environment during the measurement; that the rate of movement of the red cell which is observed may be less than that of the plasma; and that the correlation between the direction of fluid movement and the difference between colloid osmotic pressure and hydrostatic pressure is only an average one, the individual measurements showing considerable scatter.

In view of the criticisms of the results obtained by Landis's technique it seemed desirable to test Starling's hypothesis independently. This can be done by studying the rate of development of oedema in perfused preparations. It was shown by Danielli (1940) that if Starling's hypothesis is correct, and if the flow in the capillaries obeys Poiseuille's equation, then the total rate of development of oedema, R, in the preparation is related to the colloid osmotic pressure, p, of the perfusate by the equation

$$R = xp^2 - yp + \bar{z}, \qquad (1)$$

where x, y, and $\bar{z}$ are constants. Consequently, if it can be shown that this equation is obeyed, and that flow in the capillaries obeys Poiseuille's equation, then it follows that Starling's hypothesis must be obeyed by the perfused preparation.

The first attempts to discover whether Poiseuille's equation describes the rate of flow in the blood capillaries was made by Landis (1933). He studied by micro-cinematophotography the rate at which an opaque substance, when suddenly added to the blood stream, entered the individual capillaries of a network. According to Poiseuille's equation the rate should be proportional to the fourth power of the radius of the capillary. Landis found that Poiseuille's equation is obeyed to a very rough approximation. His technique was intrinsically inaccurate, measured by the standards which are necessary for testing Poiseuille's equation in such a system. For example, it would be easy to make an error of the order of 10 % in the measurement of the diameter of the relatively narrow capillary. Owing to the fourth power law, this will involve an error of the order of 50 % in calculating the rate of flow through the capillary. Another, and more satisfactory approach, is the study of the hydrodynamics of flow through leaky tubes, such as was embarked upon by Turton (1941) in a very valuable paper. In the derivation of equation (1) it was assumed that the gradient of hydrostatic pressure along a capillary of uniform bore will be linear in the steady state. If an appreciable amount of fluid is lost from the bore of the capillary owing to filtration, the pressure gradient will not be linear and equation (1) cannot be used to test Starling's hypothesis. It is therefore necessary to know whether, as a result of filtration through the capillary wall, there is normally a significant deviation from a linear gradient of hydrostatic pressure. Turton obtained a solution for the special case in which the colloid osmotic pressure of the fluid outside the capillary is constant. This particular condition is not likely to be realized in practice with exactness, but it is particularly useful because it enables us to consider the extreme cases, when the colloid osmotic pressure outside the capillary is zero (i.e. the capillary is completely impermeable to colloid), and when the colloid osmotic pressure outside the capillary is equal to that inside the capillary (i.e. the capillary is completely permeable to colloid). The former case involves smaller deviations from Poiseuille's equation than are likely to occur in practice, and the latter case involves much larger deviations than are likely to occur in practice. Consequently, if Poiseuille's equation is a satisfactory approximation for these two extremes, it will also be a satisfactory approximation for intermediate values of colloid osmotic pressure, and for cases where small variations occur in the peri-capillary colloid osmotic pressure.

Let x = distance along a capillary, measured from the arteriolar end,
r_c = radius of capillary,
l_c = length of capillary,
η = viscosity of fluid inside the capillary,
k = permeability of capillary wall,
P_a = hydrostatic pressure at $x=0$, i.e. at the arteriolar end of the capillary,
P_v = hydrostatic pressure inside the capillary at $x=l_c$, i.e. at the venular end of the capillary,
p' = colloid osmotic pressure outside the capillary,
$p_0 = p - p'$ = difference between colloid osmotic pressures inside and outside the capillary.

Turton's (1941) solution of the problem is as follows: if (as can always be tested experimentally) the amount of fluid lost by filtration through the capillary wall is a small fraction, say less than 0·1, of the flow through the capillary, then the colloid osmotic pressure inside the capillary will be constant throughout the length of the capillary. Then also

$$p_0 = p - p' \doteq \text{constant}.$$

If the capillary wall is freely permeable to protein, $p = p'$ and $p_0 = 0$. If the wall is completely impermeable to protein, $p' = 0$ and $p_0 = p$. The accuracy of the assumption that the leak from the capillaries is much less than the flow can also be tested by calculation, for in no case will a leak be greater than that given by the distribution of hydrostatic pressure for the case in which Poiseuille's equation holds, and the flow will never be less than that for the case in which Poiseuille's equation holds. So that we can calculate maximum values for the leak, and minimum values for the flow, using Poiseuille's equation:

Leak from capillary

$$= \left[\frac{P_a + P_v}{2} - p_0\right] 2\pi r_c l_c k \text{ per sec.} \qquad (2)$$

Flow through capillary

$$= \frac{(P_a + P_v)\,\pi r_c^4}{8\eta l_c} \text{ per sec.} \qquad (3)$$

Let $\alpha^2 = 16k\eta/r_0^3$. Then, according to Turton, if $\alpha^2 l_c^2$ can be neglected compared with unity, Poiseuille's equation applies accurately, the pressure gradient down the capillary is linear in the steady state, and the pressure, P, at any point in the capillary is given by

$$P = \frac{P_a(l_c - x) + P_v(x)}{l_c}. \qquad (4)$$

At the point half-way along the capillary (i.e. where $x = \frac{1}{2}l_c$)

$$P = \tfrac{1}{2}(P_a + P_v). \qquad (5)$$

If $\alpha^2 l_c^2$ cannot be neglected compared with unity, but $\alpha^2 r_c^2/4$ is negligibly small compared with unity,

$$\text{Leakage} = \frac{2\pi k r_c}{\alpha} \frac{(P_a + P_v - 2p_0)\,(\cosh \alpha l_c - 1)}{\sinh \alpha l_c}, \qquad (6)$$

$$P = p_0 + \frac{(P_a - p_0)\sinh \alpha(l_c - x) + (P_v - p_0)\sinh \alpha x}{\sinh \alpha l_c}. \qquad (7)$$

Half-way along the capillary (where $x = \frac{1}{2}l_c$),

$$P = p_0 + \frac{(P_a + P_v - 2p_0)\sinh \frac{1}{2}\alpha l_c}{\sinh \alpha l_c}. \qquad (8)$$

Equation (8) is particularly useful, for if Poiseuille's equation does not apply, the effect of the deviation

is to produce a dip in the curve of P plotted against z (Fig. 4) and the difference between the values of P at $z=\frac{1}{2}l_c$, as given by equations (5) and (8), is a measure of the size of the dip. If the dip is small, Poiseuille's equation may be used as a first approximation. If $\alpha^2 r_c^2/4$ is not negligible compared with unity, no solution is yet available.

Let us now examine the experimental results for frog capillaries to see how far Poiseuille's equation is applicable. In the experiments of Landis (1927, 1928), the radius r_c of the capillaries would hardly be less than $3{\cdot}5\mu$. The narrower the capillary, the greater the ratio surface/volume and consequently the greater the ratio leak/flow, i.e. the greater the deviation from Poiseuille's relationship. We will therefore take this minimum value of $3{\cdot}5\mu$ for r_c, since this will give us the largest possible deviation from Poiseuille's relationship. According to Landis (1934) $P_a=14{\cdot}5$ cm. H_2O, $P_v=10$ cm. H_2O, and

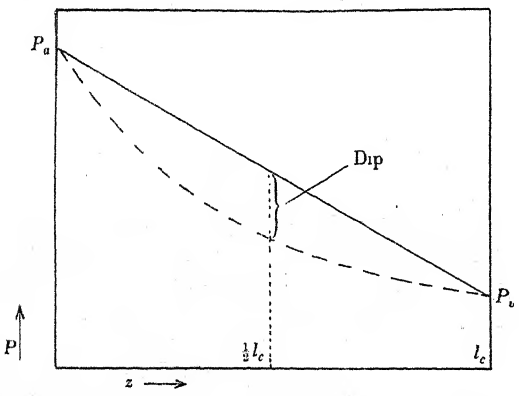


Fig. 4. Diagram of capillary blood pressure P, plotted against capillary length z. Continuous line is Poiseuille's case; broken line is case of leaky capillary, i.e. one permeable to plasma protein.

$p_0\simeq 12$ cm. H_2O: and the permeability constant k is $370\mu^3/\mu^2$/atmosphere pressure difference/minute. The viscosity of frog plasma is approximately 2×10^{-2} poise, and l_c is not likely to exceed 0·2 cm. The value of k in c.g.s. units is

$$k=\frac{370\times 10^{-12}\times 10^{8}}{76\times 13{\cdot}6\times 981}.$$

From (2) the maximum leak is

$$(12{\cdot}25-12)2\pi\times 3{\cdot}5\times 10^{-4}\times 0{\cdot}2\times 6{\cdot}08\times 10^{-10}\times 981.$$

From (8) the minimum flow is

$$\frac{(14{\cdot}5-10)\pi\times(3{\cdot}5\times 10^{-4})^4\times 981}{8\times 2\times 10^{-2}\times 0{\cdot}2}.$$

$\therefore$ Maximum leak/minimum flow = 0·005.

Hence, for the case studied by Landis, the assumption that the leak is always much less than the flow is justified, and the colloid osmotic pressure inside the capillary cannot vary significantly with z, i.e. will be constant along the whole length of the capillary. In Danielli's (1940) experiments the capillaries were dilated ($r_c\geqslant 7\mu$), and the perfusion pressure was 14 cm. H_2O, less than half the arterial pressure for a frog. Consequently, if the permeability were unchanged, and p_0 were unchanged, in these cases the ratio leak/flow$\simeq$0·001. As, however, the capillary permeability was often higher than in the normal frog, and p_0 lower than normal, it is advisable to examine the experimental ratios. Table 1 contains two sets of such values; the upper half of the table for conditions under which an unusually high rate of filtration was observed. In some cases the ratio leak/flow was as large as 0·2, so that 20 % of the fluid passing through the capillary was lost by filtration. If no protein had accompanied this fluid across the capillary wall, there would have been a substantial gradient of colloid osmotic pressure along

Table 1. *The ratio leak/flow for Hungarian frogs perfused with Ringer's solution containing various colloids*

Perfusion fluid	Leak/flow
5 % serum	0·038
1·33 % haemoglobin (cryst.)	0·23
1·75 % acacia	0·18
3·5 % acacia	0·08
40 % ox serum	0·001–0·015
2·5 % haemoglobin (crude)	0·001–0·030
2·5 % haemoglobin (cryst.)	0·08
2·2 % ovalbumin	0·09

the capillary. Actually in the cases mentioned there was also marked permeability of the capillary wall to colloid, so that the deviation from constant capillary colloid osmotic pressure will be less than is suggested by the figures in the table. The lower half of the table contains figures for more usual rates of filtration: it will be seen that changes in the colloid osmotic pressure of the perfusion fluid cannot be serious in these cases.

The initial assumption, then, that leak is much less than flow, can be taken as correct, except for a few exceptionally high rates of oedema formation. We can now examine Turton's conditions. In the case considered above, under Landis's conditions,

$$\alpha^2 l_c^2=\frac{16k\eta}{r_c^3}l_c^2$$
$$=\frac{16\times 6{\cdot}08\times 10^{-10}\times 2\times 10^{-2}}{(3{\cdot}5\times 10^{-4})^3}\,0{\cdot}2^2=0{\cdot}172.$$

Obviously $\alpha^2 l_c^2$ is not negligible compared with unity, so that Poiseuille's equation cannot apply accurately. The question arises: can it be applied approximately? $\alpha^2 r_c^2=5{\cdot}25\times 10^{-7}$, i.e. is negligible compared with unity, so that equations (6), (7) and

(8) are applicable. From (8), we have at the mid-point of the capillary ($x = \frac{1}{2}l_c$)

$$P = p_0 + \frac{(P_a + P_v - 2p_0)\sinh 0{\cdot}207}{\sinh 0{\cdot}415}$$

$$= 0{\cdot}494(P_a + P_v) + 0{\cdot}012p_0.$$

Hence, if the capillary membrane is freely permeable to protein so that $p_0 = 0$, $P = 0{\cdot}494(P_a + P_v)$. If the membrane is completely impermeable to protein $p_0 = 12 \doteq 0{\cdot}5(P_a + P_v)$, so that here $P = 0{\cdot}5(P_a + P_v)$ to a very close approximation. But, if Poiseuille's equation applied accurately, $P = 0{\cdot}5(P_a + P_v)$. Hence the maximum value of the dip is less than 1·5 % of P, so that very little error can be introduced by using Poiseuille's equation for the normal frog's capillary.

Similarly, although Poiseuille's equation does not apply exactly in Danielli's perfused frog experiments, it does apply to a sufficient degree of accuracy. For example, where the permeability of the capillaries was of the same order as that found in Landis's experiments, the dip is less than 0·5 %, and even if the permeability were 100 times greater Poiseuille's equation would still give a reasonable approximation.

We may conclude, therefore, that with the exception of cases where the rate of filtration is very large indeed, Poiseuille's equation may be used in analysing the results obtained with frog capillary systems. If greater accuracy is required, it can be obtained by using equations (6), (7) and (8). It therefore follows that if, in perfusion experiments, equation (1) is obeyed, then Starling's ultrafiltration hypothesis gives an adequate description of the fluid movements in the capillaries and peri-capillary spaces. Danielli (1940) showed that in perfused preparations this equation is obeyed where the colloid is gum acacia, crystallized ovalbumin or crystalline haemoglobin. On the other hand, when serum or crude haemoglobin are used as a source of colloid this agreement is not found, and the discrepancy was traced to changes in capillary permeability with change in colloid concentration.

The general conclusion follows that Starling's hypothesis gives a description of fluid movements in capillary systems which is accurate within the limits of our experimental errors. There are no experiments which are incompatible with Starling's hypothesis. The general conclusion that Starling's hypothesis is correct should enable us to proceed with more confidence in analysing the mechanisms of exchange of substances between the blood stream and the tissues. Krogh (1929) has argued in support of the thesis that nutritive substances reached the tissues mainly by simple diffusion through the walls of the capillaries, i.e. by direct passage through the endothelial cells. But the bulk process of filtration of Starling's mechanism provides a second, and in many cases probably much more important, mechanism of exchange. Danielli & Davson (1941) studied the penetration of sugars from the blood stream into the peri-capillary spaces, using a perfusion pressure so low that the hydrostatic pressure in the capillaries was always less than the colloid osmotic pressure of the perfusate. Under these conditions, supply of fluid to the peri-capillary spaces by the Starling mechanism is impossible, except insofar as the capillary walls are permeable to the colloid. Under these conditions sugars reach the peri-capillary spaces very slowly, but even here, where the filtration mechanism is reduced to a minimum, the kinetics of penetration suggest that filtration is more important than diffusion, as indeed would be expected if the endothelial cells had a permeability as low as other cells whose permeability has been directly determined. On the other hand, if the perfusion pressure is raised to a value of the same order as the normal blood pressure, filtration through the capillary walls occurs rapidly and a greatly enhanced rate of penetration of glucose ensues: the glucose is, of course, carried along with all the other crystalloid constituents of the perfusate in the bulk movement of the ultrafiltrate. These observations suggest that while substances such as water, oxygen, carbon dioxide and lactate, to which cells are relatively permeable, are exchanged across the capillary wall mainly by simple diffusion, other substances such as glucose, calcium and, say, insulin, to which cells are relatively impermeable, are probably mainly exchanged by the fluid movements of Starling's mechanism.

IV. LOCAL CONTRIBUTIONS TO THE CONTROL OF THE CIRCULATION

That part of the control of blood flow through the capillaries is determined by the local conditions of the tissues is now generally accepted. Krogh, in his book (1929), has presented convincing evidence on this point. Perhaps the most striking phenomenon to the eye is that in a relatively passive tissue comparatively few of the capillaries are open at any one time and that there is an alternation between open and closed capillaries. Zweifach (1939) has obtained evidence that flow normally occurs continuously through certain of the capillaries, which he calls a.-v. capillaries, and that the alternation of activity occurs only in capillary loops branching out from the a.-v. capillaries. He finds that in a perfused preparation in which there is a high rate of oedema formation flow occurs mainly through the a.-v. capillaries, but if particulate matter is added to the perfusate the capillary loops open. He attributes this opening of the capillary loops to disturbances, presumably eddies, initiated at the junction of the capillary loop with the a.-v. capillary by the particulate matter. This we believe to be a fallacy. It is doubtful whether very fine particulate matter could initiate such eddies. It is certain that the normal flow of blood in a.-v. capillaries, especially with a pulsating pressure, will cause eddies to form at the mouths of capillaries through which no flow is

occurring. It is unlikely that these eddies could be observed except in the presence of particulate matter, the deviation of which from a stream-line path will be caused by eddies. In fact particulate matter is likely to act as an indicator of pre-existing eddies rather than as an initiator of eddies. This, however, leaves unexplained the fact that particulate matter at one and the same time reduces the rate of development of oedema and causes the capillary loops to open. Krogh (1929) has suggested that a capillary loop opens when there is a sufficient lack of some metabolite, or a sufficient accumulation of some product of metabolism. If this suggestion is correct, the explanation of the action of particulate matter may be as follows: when the perfusate contains no particulate matter the rate of filtration is high, and though only a few capillaries may be open, there is a large flux of fluid through the peri-capillary spaces which is sufficient to maintain the condition of the tissues. When particulate matter is added to the perfusate it reduces the capillary permeability by blocking the pores, and so reduces the flux of fluid through the peri-capillary spaces. This leads to a deterioration in the condition of the tissue, due for instance to lack of oxygen, the response to which is the opening of capillary loops. It is of interest that this opening of capillary loops must reduce the net hydrostatic pressure in the local circulation and thus lead to a further reduction in rate of filtration in spite of the increased area available for filtration, although the rate of flow of blood through the tissue is enhanced. It is possible that part of the remarkable efficacy of particulate matter, for example, platelets, in reducing abnormally high capillary permeability is due to this decline in capillary blood pressure caused by the opening of capillary loops, and that only part of the action of the particles is due to the blocking of pores.

As was pointed out in § II a given flow of blood through the tissues may be achieved through a small number of capillaries which are widely dilated, or through a large number of capillaries which are relatively constricted. The amount of filtration in the two cases may be quite different. Thus a tissue may locally adapt its blood supply to meet its local needs without change in the arterial or the arteriolar conditions. For example, if there is a necessity mainly for the supply of oxygen and removal of carbon dioxide and lactate, this may be efficiently achieved by a relatively small number of alternating dilated capillaries, since the endothelial cells are readily permeable to these substances. But if there is a need for a large flow of fluid to the peri-capillary spaces or for the supply of, say, glucose in quantity, this result may be efficiently achieved with the same flow of blood as before through a relatively large number of relatively constricted capillaries. A sudden call for oxygen may be met by general dilatation of all the capillaries, when the rate of flow of blood will be greatly increased, and yet the rate of filtration may even fall as a result of the drop in capillary blood pressure caused by the decreased resistance to flow.

There has been considerable controversy over the interpretation of results obtained with sympathectomized animals, which seems to be due to failure to realize that the fluid exchanges across the capillary wall occur mainly by filtration, and that dilatation of the capillaries may under some circumstances cause changes in the capillary hydrostatic pressure sufficiently large to mask effects due to arteriolar dilatation. In tissues where the sympathetic normally causes a constriction of capillaries, sympathectomy will cause a dilatation of capillaries and an increased rate of blood flow. If the sympathectomy causes a greater decrease in the arteriolar resistance to blood flow than in the capillary resistance, then the hydrostatic pressure in the capillaries will rise and the rate of penetration of substances from the blood through the capillary wall to the peri-capillary fluid will be increased: at the same time there may well be a decrease in the rate at which fluid is recovered from the peri-capillary fluid into the blood stream. But if the sympathectomy causes a smaller decrease in the resistance due to the arterioles than in the resistance due to the capillaries, the capillary blood pressure will fall; there will in consequence be a decline in the rate of filtration through the capillary wall, if this fall in blood pressure is sufficient in magnitude to overcome the effect due to the increased area of capillary wall available for filtration. Such a decrease in the rate of passage from the blood stream to the peri-capillary spaces will be accompanied by an increase in the rate at which fluids such as saline and solutions of dyes are recovered from the peri-capillary fluid. Thus, where the main movement of fluid is due to the Starling mechanism, a decrease in rate of penetration in one direction in a given capillary bed is likely to be accompanied by an increase in rate of penetration in the opposite direction. This phenomenon is quite different to that usually found where penetration of membranes occurs as a result of a pure diffusion process: in this case an increase in permeability in one direction is bound to be accompanied by an increase in permeability in the reverse direction. When these differences between the processes of membrane penetration by diffusion and by bulk filtration are borne in mind and related to the hydrodynamic consequences of capillary dilatation, the conflicts of opinion between students of sympathectomized animals which have been reviewed by Gellhorn (1929) and more recently by Engel (1941) largely disappear.

V. SUMMARY

(1) Outside the single layer of endothelial cells composing the basic unit of the vertebrate blood capillary is a perithelium of connective tissue fibres, associated with which are connective tissue cells, contractile (Rouget) cells, etc.
(2) Constriction of the lumen of a capillary may occur (*a*) by the intrusion of a few endothelial cells into the

lumen, (*b*) by a more general activity of a length of endothelium, or (*c*) by the action of extra capillary cells, in which case folding of the capillary may ensue. (3) The rate of filtration through the capillary walls is controlled by (*a*) mechanical factors, (*b*) plasma and hormone factors, (*c*) local tonus and nervous control, (*d*) dietary factors. (4) The conclusion that there is a gradient of permeability to certain dyes is due to a misunderstanding of the experimental observations: the distribution of dye in such cases is precisely that predictable by Starling's ultrafiltration hypothesis for the case of a capillary slightly permeable to colloid. (5) Poiseuille's equation applies fairly accurately to the flow in blood capillaries, and the rate of development of oedema in perfused preparations obeys an equation based on Starling's hypothesis and on Poiseuille's equation. (6) Most substances reach the tissue cells as a result of filtration through the pores in the intracellular cement. Only a few substances, such as oxygen, pass with readiness through the endothelial cells themselves. (7) The capillary network may undergo changes which, with a given systemic blood flow, permit of independent variation (*a*) in the rate of blood flow through a tissue, (*b*) in the rate of filtration through the walls of the capillaries.

VI. REFERENCES

Barbieri, D. (1934): *Arch. ital. Anat. Embriol.* **32**, 602.
Bargmann, W. (1930): *Z. Zellforsch.* **11**, 1.
Barksdale, J. S. (1926): *J. Lab. Clin. Med.* **11**, 1053.
Bedson, S. P. (1922): *J. Path. Bact.* **25**, 102.
Beecher, H. K. (1936): *Skand. Arch. Physiol.* **73**, 1, 123.
Benninghoff, A. (1926): *Z. Zellforsch.* **4**, 125. — (1930): In Moellendorff's *Handb. d. mikr. Anat. d. Mensch.* **6**, 1, 18.
Bensley, R. R. & Vimtrup, Bj. (1928): *Anat. Rec.* **39**, 37.
Bielschowsky, M. & Green, H. N. (1943). *Lancet*, **245**, 153.
Busch, E. (1929): *Acta Path. Skand.* **2**, 1 (Suppl.).
Chambers, R. & Zweifach, B. W. (1940): *J. Cell. Comp. Physiol.* **15**, 255.
Clark, E. R. & Clark, E. L. (1935*a*): *Amer. J. Anat.* **35**, 239. — (1925*b*): **35**, 265. — (1939): **64**, 251.
Clark, E. R., Clark, E. L. & Abell, R. G. (1933): *Anat. Rec.* **55** (Suppl.), 50.
Clark, E. R., Hitchler, W. J., Kirby-Smith, H. T., Rex, R. O. & Smith, J. H. (1931): *Anat. Rec.* **50**, 129.
Clark, G. A. (1934): *J. Physiol.* **80**, 429.
Dale, H. H. & Richards, A. N. (1919): *J. Physiol.* **52**, 110.
Danielli, J. F. (1940): *J. Physiol.* **98**, 109.
Danielli, J. F. & Davson, H. (1941): *J. Physiol.* **100**, 246.
Defrise, A. (1929): *Monit. zool. ital.* **40**, 391.
Doi, Y. (1920): *J. Physiol.* **54**, 227.
Drinker, C. K. (1927): *J. Physiol.* **63**, 249. — (1942): *The Lymphatic System.* Stanford, California.
Drinker, C. K. & Field, M. E. (1933): *Lymphatics, Lymph and Tissue Fluid.* Baltimore.
Drinker, C. K. & Joffey, J. M. (1942): *Lymphatics and Lymphatic Tissues.* Oxford.
Duran-Reynals, F. (1939): *Yale J. Biol. Med.* **11**, 601.
Ebbeke, U. (1923): *Klin. Wschr.* **2**, 1341.
Engel, D. (1941): *J. Physiol.* **99**, 161.
Field, M. E. (1935): *Skand. Arch. Physiol.* **72**, 175.
Fieschi, A. & Storti, E. (1929): *Boll. Soc. med.-chir. Pavia*, **43**, 523.
Fleisch, A. (1937): *Pflüg. Arch. ges. Physiol.* **239**, 345.
Fleisch, A. & Weger, P. (1937): *Pflüg. Arch. ges. Physiol.* **239**, 354, 362, 476.
Florey, H. W. & Carleton, H. M. (1927): *Proc. Roy. Soc.* B, **100**, 23.
Freed, S. C. & Lindner, E. (1941): *Amer. J. Physiol.* **134**, 258.
Gabbe, E. (1926): *Z. ges. exp. Med.* **51**, 728.
Gellhorn, E. (1929): *Das Permeabilitätsproblem.* Berlin.
Green, H. N. (1943): *Lancet*, **245**, 147.
Golubew, A. (1869): *Arch. mikr. Anat.* **5**, 49.
Harris, E. & Marvin, H. M. (1927): *Heart*, **14**, 135.
Hartman, F. A., Evans, J. I., Melachowski, B. T. & Michalek, L. M. (1928): *Amer. J. Physiol.* **85**, 99.
Hartman, F. A., Evans, J. I. & Walker, H. G. (1928): *Amer. J. Physiol.* **85**, 91.
Heimberger, H. (1927*a*): *Z. Zellforsch.* **4**, 713. — (1927*b*): *Z. ges. exp. Med.* **55**, 17.
Hemingway, A. & McDowell, R. J. S. (1927): *J. Physiol.* **62**, 166.
Hooker, D. R. (1920): *Amer. J. Physiol.* **54**, 30.
Huzella, T. (1925). *Z. Zellforsch.* **2**, 558.
Hyman, C. & Chambers, R. (1943): *Endocrinology*, **32**, 310.
Jones, T. (1936): *Amer. J. Anat.* **58**, 227.
Jores, A. (1927): *Z. ges. exp. Med.* **59**, 172.
Kahn, R. H. & Pollak, F. (1931): *Pflüg. Arch. ges. Physiol.* **226**, 799.
Krogh, A. (1920): *J. Physiol.* **53**, 399. — (1922, 1924, 1929, 1937): *The Anatomy and Physiology of Capillaries.* Newhaven.
Krogh, A. & Harrop, G. A. (1921): *J. Physiol.* **54**, cxxv.
Krogh, A., Harrop, G. A. & Rehberg, P. B. (1922): *J. Physiol.* **56**, 179.
Krogh, A. & Vimtrup, Bj. (1932): In *Special Cytology*, 1, Edited by Cowdry. New York.
Landis, E. M. (1927): *Amer. J. Physiol.* **82**, 217. — (1928): **83**, 528. — (1933): **103**, 432. — (1934): *Physiol. Rev.* **14**, 404.
Leriche, R. & Policard, A. (1920): *C.R. Soc. Biol., Paris*, **20** (November).
Lewis, T. (1927): *The Blood Vessels of the Human Skin.* London.
Lewis, T. & Marvin, H. M. (1927): *Heart*, **14**, 27.
Loeschke, H. & Loeschke, E. (1934): *Z. mikr.-anat. Forsch.* **35**, 533.
McClean, D. (1930): *J. Path. Bact.* **33**, 1045.
McMaster, P. D., Hudack, S. S. & Rous, P. (1932): *J. Exp. Med.* **55**, 203.
Marchand, F. (1923): *Münch. med. Wschr.* **70**, 385.
Mayer, S. (1899): *Anat. Anz.* **16**, 180. — (1902): **21**, 442.
Menkin, V. (1940): *The Dynamics of Inflammation.* New York.
Michels, N. A. (1936): *Anat. Rec.* **65**, 99.
Moore, R. L. (1936): *Anat. Rec.* **64**, 387.
Nagel, A. (1934): *Z. Zellforsch.* **21**, 376. — (1937): *Z. mikr.-anat. Forsch.* **42**, 433.
Pfuhl, W. (1933): *Z. Zellforsch.* **20**, 390.
Plenk, H. (1927): *Ergebn. Anat.* **27**, 302. — (1930): *Anat. Anz.* **69**, 25. — (1937): *Z. mikr.-anat. Forsch.* **42**, 467.

RINEHART, J. F. (1930): *Amer. J. Path.* **6**, 525.
ROGERS, J. G. (1932): *Anat. Rec.* **54**, 1. — (1935): **63**, 193.
ROUGET, CH. (1873): *C.R. Acad. Sci., Paris*, **79**, 559. — (1879): **88**, 916.
ROUS, P., GILDING, H. P. & SMITH, F. (1930): *J. Exp. Med.* **51**, 807.
ROUS, P. & SMITH, F. (1931): *J. Exp. Med.* **53**, 219.
SALVIOLI, G. (1927): *Sperimentale*, **81**, 215.
SANDERS, A. G., EBERT, R. H. & FLOREY, H. W. (1940): *Quart. J. Exp. Physiol.* **30**, 281.
SANDISON, J. C. (1931): *Anat. Rec.* **50**, 355. — (1932): **54**, 105.
SASLOW, G. (1938): *Amer. J. Physiol.* **124**, 360.
SCHAFFER, J. (1920): *Vorlesungen über Histologie und Histogenese.* Leipzig.
SCHULZE, W. (1925): *Z. ges. Anat.* 1. *Z. Anat. EntwGesch.* **76**, 421.
SMITH, F. & DICK, M. (1932): *J. Exp. Med.* **56**, 371.
SMITH, F. & ROUS, P. (1931*a*): *J. Exp. Med.* **53**, 195. — (1931*b*): **54**, 499.
STARLING, E. H. (1895): *J. Physiol.* **19**, 312.
STEINACH, E. & KAHN, R. H. (1903): *Pflüg. Arch. ges. Physiol.* **97**, 105.
STRICKER, S. (1865): *S.B. Akad. Wiss. Wien*, Math-Nat. Kl., **52**, 379. — (1879): **74**, 313.
SWINGLE, W. W., PFIFFNER, J. J., VARS, H. M. & PARKINS, W. M. (1933): *Science*, **77**, 58. — (1934): *Amer. J. Physiol.* **107**, 259.
TANNENBERG, J. (1926): *Dtsch. med. Wschr.* **10**, 404.
TARCHANOFF, J. F. (1874): *Pflüg. Arch. ges. Physiol.* **9**, 407.
TURTON, F. J. (1941): *Phil. Mag.* **32**, 457.
URTUBEY, L. (1932): *Rev. Esp. Biol.* **1**, 25.
VENZLAFF, W. (1911): *Arch. mikr. Anat.* **77**, 377.
VIMTRUP, BJ. (1922): *Z. ges. Anat.* 1. *Z. Anat. Entw-Gesch.* **65**, 150. — (1923): **68**, 469.
VOLTERRA, M. (1925*a*): *Monit. zool. ital.* **36**, 49. — (1925*b*): *Sperimentale*, **79**, 618. — (1925*c*): *Zbl. inn. Med.* **46**, 876. — (1933): *Monit. zool. ital.* **44** (Suppl.), 71.
VOLTERRA, M. & SCHUPFER, F. (1934): *Arch. ital. Anat. Embriol.* **33**, 844.
WIND, F. (1937): *Arch. Exp. Path. Pharmak.* **186**, 161.
WOOLLARD, H. H. (1928): *Med. J. Austr.* **2**, 134.
ZIMMERMANN, K. W. (1886): *Arch. mikr. Anat.* **27**. — (1923): *Z. ges. Anat.* 1. *Z. Anat. EntwGesch.* **68**, 29.
ZWEIFACH, B. W. (1934): *Anat. Rec.* **59**, 83. — (1937): *Amer. J. Anat.* **60**, 473. — (1939): *Anat. Rec.* **73**, 475. — (1940): *Proc. Soc. Exp. Biol., N.Y.*, **44**, 124.
ZWEIFACH, B. W. & KOSSMANN, C. E. (1937): *Amer. J. Physiol.* **120**, 23.

THE GOLGI APPARATUS OF PROTOZOA

BY J. D. SMYTH, PH.D., Department of Zoology, University College, Leicester

(*Received* 11 *April* 1944)

I. INTRODUCTION

The Golgi apparatus problem in Protozoa has received much attention within the last 25 years and a great mass of conflicting data has been accumulated. The earlier work has been reviewed by King (1927) and Hill (1933), and certain aspects of the subject have been summarized briefly by MacLennan (1941). Recent work on centrifuged (Singh, 1938; Daniels, 1938) and dividing organisms (Gatenby & Singh, 1938; Gatenby & Smyth, 1940; Smyth, 1941) has added to our knowledge in this field, and an up-to-date review would seem to be of value.

In metazoan cells the Golgi apparatus is easily demonstrable after certain osmic (Weigl & Kolatchew) or silver techniques (Aoyama, Cajal, etc.) as it possesses the power of reducing these heavy metals. The invention of the ultra-centrifuge by Beams (1930) and its application to cytological problems by Beams, Weed & Pickles in 1933, and later by many others, proved beyond doubt that such an apparatus was a true cellular inclusion and not a precipitation artifact as had been claimed by earlier workers (Parat & Painlevé, 1924; Walker & Allen, 1927; Walker, 1928). The morphology of the Golgi apparatus in metazoan cells is very variable. It has been described in various cells as having the form of a network, a fibrous reticulum, a collection of rods, platelets, or disks, a hollow sphere, a group of vacuoles—just to mention a few types. It is especially large in gland cells and often exhibits notable polarity, frequently going through a series of changes in correlation with changes in cellular activity. Kirkman & Severinghaus (1938 *a*, *b*, *c*) and Bourne (1942) have recently summarized our knowledge in this field.

The Golgi techniques were first applied to Protozoa by Hirschler (1914) who demonstrated the presence of a series of crescents and vesicles which greatly resembled the metazoan Golgi apparatus in possessing the typical osmiophile and osmiophobe regions. Hirschler worked with *Monocystis ascidiae, Diplocystis phryganeae, Gregarina blattarum* and *G. polymorpha*. A few years later Gatenby & King (1923) investigated *Adelea ovata* and obtained somewhat similar results, describing osmiophile bodies which gave the typical reactions of the metazoan Golgi apparatus; these they identified as Golgi elements. Joyet-Lavergne in the same year also obtained comparable results in *Adelina demidiata*. This early work on the Sporozoa, was followed in 1924 by a paper on flagellates and ciliates by Nassonow, opening up a

line of research which has formed the real basis for most of the later work on the protozoan Golgi bodies. Using mainly the Kolatchew osmic technique he demonstrated the presence of an osmiophile cortex surrounding, or in close relation to, the contractile vacuole of *Nassula laterita*, *Lionotus folium*, *Chilomonas paramecium*, *Vorticella* sp., *Campanella umbellaria*, *Epistylis gallea*, and *Zoothamnium arbuscula*. In *Paramecium caudatum* he described long osmiophile canals leading into the thin-walled contractile vacuole. From the fact that the osmiophile material in association with the vacuole was apparently secretory, lipoidal, and had the power of reducing osmium after the standard Golgi methods, Nassonow concluded that together with the contractile vacuole it represented the homologue of the Golgi apparatus of higher forms. He believed that in Protozoa it served as a mechanism for the collection of certain cellular contents and their excretion into the vacuole.

Gatenby & King (1923) suggested that the Golgi apparatus arose in connexion with the parabasal body of the flagellum of some primitive flagellate. This theory has received much support from the results of Duboscq & Grassé (1925, etc.) They have described the parabasal body in a number of forms (see later) and have shown that it is clearly demonstrable after impregnation by Cajal or Weigl methods, and frequently has the same chromophile-chromophobe structure as the Golgi apparatus of metazoan germ cells. Moreover, they have emphasized the secretory activity of the parabasal during division, which is also a point in favour of the Golgi apparatus homology.

Since this early work of Nassonow and his contemporaries an immense amount of material has been studied by numerous investigators in every class of Protozoa. The most controversial point that has arisen in connexion with this work is the nature of the criteria for identification of the Golgi apparatus, and it is to this point that attention must first be directed.

II. TECHNIQUE

Even in the Metazoa identification of the Golgi apparatus in various forms is considered indefinite by some workers. MacLennan (1941) states: 'The selection of criteria is a critical point in the identification of Golgi material, yet even after years of work on representatives of all the major groups of animals and plants, and not withstanding periodic reviews of the field, few criteria seem to have unanimous approval.' Gatenby (1930), on the other hand, states: 'Modern workers in general have experienced no difficulty in identifying Golgi bodies.' It is possible that much of the controversy concerning the results of different workers has resulted from faulty technique or the use of impure materials, especially osmic acid which is notably difficult to obtain in a pure state.

One of the major difficulties is that some workers have applied the 'vacuome' hypothesis of Parat & Painlevé (1924) to the Protozoa. The earliest case of this is that of Joyet-Lavergne (1926), who claimed that the neutral red bodies of gregarines were identical in form and distribution with the Golgi bodies in these organisms. Many later workers have taken up this claim without further confirmation as to its validity. Hall and his associates (1929–33) have been foremost in this field and have described a 'vacuome' as being present in numerous organisms; this 'vacuome' is claimed to be osmiophile, argentiophile and stainable in neutral red. Much has been written on the 'vacuome theory', but most modern workers (Tuzet, 1931; Subramaniam & Ganapati, 1938; Kirkman & Severinghaus, 1938 *a*, *b*, *c*; Gatenby, 1938–41; Smyth, 1941–4; Bourne, 1942) believe that it may be discounted on the grounds of more recent experimental evidence. Gatenby (1941) summarizes the situation: 'A good many protozoologists have published papers purporting to describe what they have called Golgi bodies in various Protozoa. Quite often these protozoologists have lacked a proper knowledge of metazoan Golgi apparatus, and what has added to the confusion, they have not hesitated to base some of their conclusions on the now defunct neutral red 'vacuome' hypothesis of Parat. This is all the more strange in view of the masterly researches of Nassonow, in which the clearly separate existence of mitochondria and contractile vacuole cortex had been demonstrated in a large number of Protozoa. If almost any living Protozoa be immersed in a weak neutral red solution, small red vacuoles appear in the cytoplasm. Transfer these stained organisms to osmic acid, and the site of each of the red vacuoles becomes a black granule. This is not a Weigl or a Kolatchew method—it merely shows that in the vacuole of the organism, the osmic tetroxide is speedily reduced by the neutral red solution. The same thing occurs when the two solutions are mixed in a test-tube.'

The work of Daniels (1938) finally gave very convincing proof of the inadequacy of the neutral red technique. She showed that in ultra-centrifuged gregarines the Golgi bodies always moved to the centripetal end of the cell and occupied the same position relative to the other inclusions as did the Golgi material in the centrifuged metazoan cell; whereas the neutral red granules were never displaced. It is not intended to review further the vast literature on the neutral red controversy in the present article, for adequate reviews of this aspect of the subject are already available (MacBride & Hewer, 1931; Kirkman & Severinghaus, 1938 *a*, *b*, *c*; Hirsch, 1939; Gatenby, 1931).

In general the Weigl and Kolatchew techniques have been used by most workers (Nassonow, Gatenby, Singh, Brown, Beams, Hirschler, Moore, Patten, King, Smyth, *et al.*), but even these methods are liable to prove capricious. Bleaching by turpentine or hydrogen peroxide helps to increase the specificity of these methods since resistance to this is more pronounced in Golgi material; but, nevertheless, it

must be admitted that in some cases the judgement of the observer is often a determining factor. This individual judgement can usually be eliminated by the use of such techniques as Altmann's acid fuchsin after osmication. Smyth (1941) pointed out the need of a 'control' in osmication procedures and suggested that a culture of *Chilomonas* should be mixed with the organisms under investigation since many workers (Nassonow, 1924; Hall, 1930*a*; Gatenby & Smyth, 1940) have shown this flagellate to possess a well-marked osmiophile cortex to its contractile vacuole. If, then, after the Golgi techniques, the cortex of the vacuole was not blackened in *Chilomonas*, it would be almost certain that impregnation in the other organisms would also be faulty. It is significant to note, for example, that Patten & Beams (1936) failed to impregnate the vacuolar cortex in *Chilomonas*, thus making their observations on the other flagellates in the same culture open to question. The use of the ultra-centrifuge on Protozoa (Daniels, 1938; Browne, 1938) has shown how successful these methods can be in the hands of careful workers. In spite of this, many workers believe the Golgi methods are not specific for any single kind of material. Tennent, Gardiner & Smith (1931) state that a series of different lipoids extracted from echinoderm eggs give a typical Golgi reaction. MacLennon (1940) states that both lipoid and non-lipoid granules of various types in the Protozoa react typically to the Golgi techniques. Hoerr (1936) suggested that osmium reduced in the tissues might migrate and be secondarily absorbed by cellular structures other than Golgi bodies.

One of the typical characteristics of the metazoan Golgi apparatus is that it can be impregnated by the silver methods of Cajal and Da Fano. Such methods seldom are successful in Protozoa, though some workers (Richardson & Horning, 1931; Duboscq & Grassé, 1925; Hall, 1929–31) have claimed to have impregnated Golgi bodies by those methods, though in general their results have not been confirmed by other workers. Browne (1938) states: 'One of the greatest stumbling blocks in protozoan cytology is the failure of the silver methods to demonstrate the Golgi apparatus.' Since the silver methods are notably capricious with invertebrate material, too much emphasis must not be laid on their failure in the Protozoa, though some workers (Brown, Hall *et al.*) believe that it is an essential criterion for identification of Golgi material in the Protozoa.

The morphology of the protozoan Golgi material, like that of the metazoan cell, varies very strikingly, and is claimed to be represented by granules, crescents, spheres, parabasal bodies, and osmiophile cortices in different species. In many organisms it varies during the life cycle of the individual, not only in size and distribution but also in osmiophility (MacLennan, 1936, 1937). Before discussing the problem further a summary of the work that has been carried out in the four classes of the Protozoa will be given.

III. OCCURRENCE

1. *Flagellata*

Since the flagellates are believed by many (Gatenby, 1941; Hyman, 1940) to be the most primitive class of the Protozoa it is well to discuss the problem first in this group. Gatenby as early as 1923 put forward the view that 'the Golgi bodies probably arose in connexion with the terminal bead of the flagellum of some primitive flagellate', and suggested that it may have been concerned with the storage of lipoid material necessary for the metabolism of the locomotor organ. Duboscq & Grassé carried out most of the early work in this field and showed that in many cases the parabasals could be impregnated by the osmic or silver Golgi techniques of Weigl or Cajal. In *Holomastigotes elongatum* (1925*a*) they described the parabasal in the young stages as consisting of a chromophile cap surrounding a chromophobe medulla. During growth the parabasals divide so that in the adult there are numerous bodies present. In *Pyrsonympha vertens* the parabasals are described by Duboscq & Grassé (1925*b*) as typical curved osmiophilic dictyosomes applied to the nuclear membrane as in many metazoan cells. They have also shown that in many flagellates the parabasal, like the Golgi apparatus of higher forms, is concerned in secretion. In *Trichomonas batrachorum*, *Tetramastix bufonis* and *Trichonympha chattoni* droplets break off from the chormophobe substance of the parabasal apparatus, apparently becoming dissolved in the cytoplasm. In *Jaenia annectans* (Duboscq & Grassé, 1928) the secretory action is particularly evident during division. In the dinoflagellate *Polykrikos schwartzi* Chatton & Grassé (1929) described the osmiophilic parabasals as three long streamer-like structures which they suggest probably secrete the oblong osmiophilic vesicles also present in the cytoplasm. Brown (1930*a*) put forward some evidence against the parabasal hypothesis by showing that in the hypermastigote flagellate *Microjienia*, there are presumed Golgi bodies present as crescents, rings or spheres distinct from the two parabasals. Hirschler (1932) investigated the cytology of *Bodo lacertae*, *Trypanoplasma helicis* and *Lophomonas blattarum*, and came to the conclusion that the parabasal of *Bodo* represents the Golgi apparatus, while that of *Trypanoplasma* probably represents the mitochondria; with regard to *Lophomonas* he came to no conclusion.

In *Euglena* sp. Grassé (1925, 1926) and Duboscq & Grassé (1933) suggested that the stigma represented the parabasal apparatus (and therefore the Golgi apparatus according to their views) of other flagellates. In support of this they pointed out that it reacts similarly to fixatives, being slowly impregnated with osmium or silver; it is made up of a chromophobe (proteid) basis, in which lie scattered chromophile (lipoid) bodies; it multiplies by binary

fission during cell division and is closely related to the blepharoplast. Patten & Beams (1936) have shown that although the stigma is readily osmicated the blackening is easily removed by bleaching, which suggests that it cannot be true Golgi material. Several workers have described the Golgi bodies in *Euglena* as being distributed throughout the cytoplasm. Thus Brown (1930*a*) described osmiophilic spheres with black borders and clear centres which were scattered in the cytoplasm, though he mentions that they are more abundant in the region of the nucleus. Hall (1931*a*) described similar bodies and identified them as 'vacuome' since they appeared after neutral red staining; Baker (1933) figures Golgi bodies similar to those described by Brown, but shows the neutral red bodies, i.e. 'vacuome', as separate entities.

Sigot (1931) was the first to show that these workers had not obtained complete impregnation of *Euglena*, and that there were 'plaquettes osmiophiles' associated with the reservoir. Gatenby & Singh (1938), on re-examining the material of Patten & Beams (1936), found that in *E. viridis* the osmiophile material lies at the lower end of the reservoir, where it forms a separate vacuole. In *E. agilis*, however, the osmiophile material is intimately related to the wall of the reservoir. In addition, various conditions of osmiophile accessory contractile vacuoles were described; and a swollen osmiophile vesicle is found ventral to the contractile vacuole or in direct contact with it. They believed the osmiophile material represented Golgi substance.

In *Chilomonas paramecium* Nassonow (1924) described a thick osmiophile cortex to the contractile vacuole. This organism was studied in greater detail by Gatenby & Smyth (1940) who confirmed Nassonow's results and showed that during division the osmiophile (Golgi) cortex separated evenly into two parts, one daughter half going to each individual. In a small proportion of dividing stages (3 %) the osmiophile ring passed over completely to one daughter monad. They regarded the osmiophile cortex as representing Golgi material. Scattered osmiophile granules were also described in the protoplasm. Hall (1930*a*), in the same organism, described osmiophilic globules blackened in 85 % of the specimens, contractile vacuole wall in 54 %, and the globules alone impregnated in 25 %. He also described two 'blackened vacuoles', which Gatenby & Smyth (1940) have identified as pyrenoids. Since 'the osmiophilic globules of *Chilomonas* are more consistently impregnated than the contractile vacuole' he believed them to represent the Golgi apparatus.

In *Copromonas subtilis* Gatenby & Singh (1938) described the Golgi material as consisting of granules which they believe are associated with the osmoregulatory mechanism of the cell. This material may form a thick cortex around the reservoir or it may be completely detached from it. A definite dictyokinesis is described as taking place, the osmiophile material moving slightly down the cell and splitting into subequal groups. This material was found to persist throughout conjugation and encystment, even when a reservoir could not be found.

Hall (1931*a*), using his usual 'vacuome' technique, in giving an account of the cytoplasmic inclusions of *Rhabdomonas* (*Menoidium*), described the Golgi bodies as scattered throughout the cell. Patten & Beams (1936) described similar globules in this organism, but did not identify them as true Golgi material. Smyth (1944) described the Golgi apparatus of *Rhabdomonas* as being represented by a spherical mass of osmiophile material lying below the reservoir; a canal with osmiophilic walls connected it to the reservoir. During division the canal disappeared, the mass became closely applied to the reservoir, elongated and finally separated into two distinct portions. In a very small number of cases all the material was transferred to one of the daughter cells formed by the division. In all the specimens of *Rhabdomonas* described diffuse osmiophile granules were present in the cytoplasm.

In the euglenoid flagellate *Astasia Harrisii* the same writer (1943–4) described a similar type of osmiophile complex, though in this case the 'canal' was longer than in *Rhabdomonas* and frequently consisted of a series of osmiophilic vacuoles placed end to end. It was believed that the osmiophile material represented the Golgi apparatus and that the whole structure—both osmiophilic mass and canal—was concerned with the excretion or secretion of some material into the reservoir.

Only one member of the Choanoflagellata has been investigated. Saedeleer (1930) described the Golgi apparatus in a solitary form as consisting of a small ring of osmiophile material lying at the base of the flagellum below the basal granule.

Although, as was mentioned earlier, the 'vacuome' (neutral red) technique of Parat and his followers is now considered by most modern workers as obsolete, it is worth while, in order to complete the summary of our knowledge of this class, to note that (including some of the types already mentioned) a 'vacuome' has been described in the following groups: Chrysodmonadida (Hall, 1930*a*); Cryptomonadida (Hall, 1930*a*); Dinoflagellata (Chatton & Grassé, 1929); Euglenida (Grassé, 1925; Hall, 1929, 1931*a*); Protomastigida (Hirschler, 1927; Lwoff & Lwoff, 1929; Phytomonadida (Hall & Nigrelli, 1930); Hypermastigida (Hirschler, 1927).

2. *Rhizopoda*

There is much confusion as regards the identification of Golgi material in this group. Brown (1930*b*) claimed to have demonstrated two types of Golgi apparatus in *Amoeba proteus* using the Weigl osmic technique. One type consisted of granules which stain darkly in osmic acid, and the other of spherules with osmiophile black rims. No Golgi material was impregnated in association with the contractile

vacuole; he claimed, however, to have shown the presence of small granules near it. He suggested that the Golgi spherules are formed from the growth and swelling of the granular type. Mast & Doyle (1935 *a*, *b*) published an exhaustive account of the cytology of *A. proteus* '*Y*' and *A. proteus* '*X*' (*A. dubia*) including the results of ultra-centrifuging. They described refractive bodies in the cytoplasm which stained with neutral red; the outer layer of these bodies gave positive Golgi tests, and these workers concluded from this that they represent the Golgi bodies of *Amoeba*, They also figure vacuole refractive bodies which stain rapidly in neutral red—more so than the refractive bodies mentioned above; these refractive vacuoles are easily stainable in osmic acid and do not readily bleach in turpentine, hydrogen peroxide or turpentine water. The refractive bodies (Golgi bodies) originate and develop in the cytoplasm from substances obtained from the vacuole refractive bodies, and the cytoplasmic crystals. They disintegrate during starvation and function in the normal individual as reserve food. Singh (1938) repeated most of the work of Mast & Doyle on *A. proteus* '*V*' and obtained somewhat different results. The granules described by Brown (1930*b*) he claims to be mitochondria as they could be bleached by hydrogen peroxide and stained in acid fuchsin or Janus B green. He could not observe relationship between mitochondria and fat, or mitochondria and nutrient spheres, and never observed the transformation of granules into spherules with dark rims, as claimed by Brown. These spherules are described by Singh as being sudanophile fat since they stain in Sudan IV or Nile blue. After centrifuging they occupied the centrifugal end of the cell. The refractive bodies of Mast & Doyle Singh calls nutrient spheres since they give the reactions of glycogen; he was never able to stain the outer portions of these by any of the Golgi techniques. He concludes that there is no structure in *A. proteus* homologous with the metazoan Golgi apparatus.

Hall (1930*b*), using the Kolatchew osmic technique, reported small globules scattered in the cytoplasm of *Trichamoeba* which resisted bleaching in turpentine or hydrogen peroxide. These he described as adhering closely to the contractile vacuole in a few instances. He states further that these Golgi bodies are identical in size and distribution with the neutral red ('vacuome') bodies. Causey (1925), using smears fixed in osmium vapour and stained with Regaud iron haematoxylin, described the Golgi elements in *Entamoeba gingivalis* as arising from food vacuoles. When these disappeared some of the vacuolar walls thickened to form deeply staining crescents which later became twisted and tangled to give a net-like appearance. In *E. blattae* Hirschler (1927), using both silver and osmic methods, described the Golgi bodies as being similar to those found in sporozoans, that is, scattered spheres or crescents with typical osmiophile and osmiophobe regions.

Hall & Loefer (1930), using the Kolatchew osmic and Da Fano silver methods, described the Golgi in *Euglypha alveolata* material as being represented by diffuse globules, in general limited in distribution to the alveolar zone. The wall of the contractile vacuole was never blackened, but slight impregnation of the food vacuoles was observed. Similar globules were also visible in the encysted stages. Hall applied the neutral red technique and identified the scattered granules as being also identical with the 'vacuome'. Similar results were obtained by Hall & Nigrelli (1930) for *Arcella*. Lamont (quoted by Gatenby, 1938), using the ultra-centrifuge and the standard osmic techniques, was unable to demonstrate the presence of any Golgi material in *Nebela collaris* (*pro parte*, Leidy).

3. *Sporozoa*

Hirschler (1914) was the first worker to describe a Golgi apparatus in the Sporozoa. He described bodies in *Monocystis ascidiae* as having the form of a series of rings or crescents which greatly resembled the metazoan Golgi bodies. He showed the mitochondria as separate inclusions. The same worker later (1924) described similar bodies in *M. agilis*. In *Gregarina polymorpha* and *G. blattarum*, which he mentions in the same paper, he was unable to distinguish between mitochondria and Golgi material and believed it represented a primitive condition of these inclusions in which the Golgi apparatus and mitochondria had not yet become differentiated. Joyet-Lavergne (1926), however, does not agree with Hirschler, and states that the Golgi bodies and mitochondria may be differentiated by careful fixation and staining. In the same genus (i.e. *Gregarina*) Daniels (1938) described the Golgi bodies as being divided into two main categories: (*a*) small granules and short rods which in the centrifuged animal are seen among the fatty globules at the extreme centripetal end of the cell, and (*b*) larger bodies, dictyosomes, and irregular rings and rods, having a specific gravity slightly greater than the (*a*) type. On staining intravitally in neutral red she described scattered globules present; even after centrifuging 'red globules appeared as before, and irregularly throughout the cytoplasm, apparently in no relation to the other inclusions, which were banded'.

Gatenby & King (1923) demonstrated a number of crescentic rods or dictyosomes in the adult trophozoite of *Adelea ovata* which they believed represented Golgi bodies. The Golgi apparatus was usually juxta-nuclear in position in the young sporozoites and spread throughout the cell during ontogony. Tuzet (1931) described the Golgi bodies of *Gonospora duboscqui* as being similar in form and behaviour. Joyet-Lavergne (1924) found that in the sporozoites and microgametes of *Aggregata eberthi* the Golgi bodies agree with the typical sporozoan type of Hirschler in form, and are very closely connected with the nucleus. In the young schizoites, the Golgi

bodies are arranged in a crescent to which they are connected by a 'siderophile axis'. At a later stage this crescent straightens out to form a rod which splits into two—one-half lying on each side of the axis; this latter stage is compared by Duboscq & Grassé (1927) to the parabasal of *Pseudotrichonympha*. In *Nina gracilis* (a gregarine) he states that the Golgi material occupied a similar position. With neutral red he describes the Golgi bodies as being the first to colour in the form of 'petits arcs ou granules'. King (1926), using mainly the Weigl method, identified the Golgi bodies in the plasmodium stage of *Haplosporidium* as scattered masses throughout the cytoplasm. In the sporoblast the Golgi apparatus is described as being large and easy to demonstrate, embracing the nucleus first as an irregular mass, but as the sporoblast develops into a spore, it moves away from this position to the end of the cell. In its final position it formed a round black mass just under the spore cap. The only haemosporidian investigated by Golgi methods is *Plasmodium praecox* in which Cowdry & Scott (1928) described osmiophilic globules distributed in the cytoplasm. These globules were often seen as loose aggregates which could fuse into filaments and nets.

4. *Ciliata*

Nassonow (1924–5) was the first to describe a Golgi apparatus in the Ciliata. Using mainly the Kolatchew technique he demonstrated osmiophile material associated with the contractile vacuolar complex in *Paramecium caudatum*, *Lionotus folium*, *Nassula laterita*, *Dogielella* sp., *Chilodon* sp., *Campanella umbellaria*, *Epistylis gallea*, *Zoothamnium arbuscula* and *Vorticella* sp. He proposed that this osmiophile material *together with* the contractile vacuole was the protozoan homologue of the Golgi apparatus in metazoan cells, and that the whole formed a secretory mechanism. Many workers have confirmed his findings in these organisms, but his proposed homology has been the subject of some dispute. In *Paramecium caudatum* Nassonow described the vacuolar canals as having thin osmiophilic walls. The distal portion of these canals is surrounded by a specially differentiated plasma, and according to Nassonow a hypertonic liquid is secreted from this area into the vacuoles. Prior to division the whole vacuolar apparatus is said to divide and the new vacuolar systems arise by the multiplication of the osmiophilic canals. Nassonow's description of the morphology of the osmiophile material in *Paramecium* has been confirmed by Von Gelei (1928) and by Smyth (1941). Dunihue (1931) stated that in the osmication of this organism the neutral red granules were impregnated first and represent the 'vacuome'. R. King (1935) described a 'specialized' excretory protoplasm surrounding the vacuolar canals of *P. multimicronucleata*, but denied Nassonow's homology.

In *Lionotus folium* Nassonow described an osmiophile mass in association with the contractile vacuole of this organism. Smyth (1941), working on the same genus (but probably a different species), describes Golgi bodies with osmiophilic cortices and osmiophobic centres extending laterally along the whole length of the body. Diffuse scattered osmiophile globules are also found in the cytoplasm. The Golgi bodies divide by simple binary fission. In *Nassula* Nassonow states that the osmiophile substance is present as a cortex to the contractile vacuole which during systole stains as a solid mass of material. In *Dogielella*, however, he described the material as forming rings around the vacuoles, like 'rings around the planet Saturn'. When the vacuole contracted the ring remained unchanged. In *Chilodon*, Nassonow showed that each of the contractile vacuoles was surrounded by a permanent osmiophilic cortex. Numerous small vacuoles are described as appearing in the substances of the ring after systole; this stage he names the 'bound secretion stage'. These small vacuoles run together to form a large vacuole enclosed by the ring; this he calls the 'free secretion stage'. Smyth (1941), in *Chilodon*, described as many as six contractile vacuoles, each with an osmiophile (Golgi) cortex, although three was the usual number. The cortex to each vacuole he described as having a small break in it so that a complete ring is not obtained. According to Nassonow (1924) a thick cortex is also present in *Campanella umbellaria*. After systole small vacuoles are formed in the substance of the osmiophile material—a condition Nassonow believed to be pathogenic. Fauré-Fremet (1925), however, confirmed these results and was able to follow out the same process *intra-vitam*, and believed it to be normal in *Campanella*. Haye (1930) confirmed the presence of an osmiophile cortex around the vacuole.

Similar results were obtained by Nassonow in other peritrichs—*Vorticella* sp., *Zoothamnium arbuscula*, and *Epistylis gallea*. Fauré-Fremet (1925), using the usual Golgi methods, confirmed these observations and stated that during systole the cortex collapses and a new vacuole is formed by the fusion of small vesicles secreted within the wall. Hall & Dunihue (1931), using the neutral red technique, described the 'vacuome' as scattered osmiophile bodies in *Vorticella convallaria*, *V. microstoma* and *V. campanula*. These globules, they stated, were stainable in neutral red and could be distinguished from mitochondria by their negative reaction towards Janus green. Gatenby (1941), in *Vorticella* sp., showed that a very thick osmiophile (Golgi) cortex was present in all the species examined. During division the contractile vacuole with its osmiophile membrane is carried over completely to one daughter organism, whereas a new vacuole is formed in the other. This vacuole at first has no osmiophile cortex, but gradually it becomes formed. Scattered osmiophile granules are also present in the cytoplasm, and this writer suggests these may take part in the formation of the

new osmiophile cortex after division. Very convincing microphotographs are shown in support of this statement. Gatenby also described a colony of *Epistyois leucoa* in which an osmiophile cortex was shown to be present in every organism, although the scattered globules present in *Vorticella* were apparently absent in this form. In the Hypotricha very few forms have been investigated. Hall (1931*b*) described the 'vacuome' in *Stylonychia* as being represented by scattered globules which were stainable by Weigl, and Kolatchew osmic methods or Da Fano and Cajal silver methods, and neutral red. Smyth (1941) believes that no true Golgi material is present in this ciliate. The only structures which he blackened after Weigl treatment were those visible in the living condition, namely, the accumulation of granules at the posterior end. In *Euplotes* sp., according to R. King (1933), the Golgi bodies are represented by small accessory osmiophilic vacuoles associated with the contractile vacuole. These arise from a very large number of canals which radiate from the vicinity of the vacuoles. These canals are also osmiophilic in nature. Smyth (1941) concluded that the Golgi apparatus was absent in this organism, for no granules were present, the osmium tetroxide being reduced to some extent by the food boli which became darkened, but never blackened.

In the large ciliate *Spirostomum ambigium* Hirschler (1924) described only one type of lipoid inclusion which he concluded represented both mitochondria and Golgi bodies. Browne (1938), in the same organism, demonstrated round Golgi bodies lying on the endoplasmic threads of the protoplasm and in the ectoplasm. Some of the bodies had a heavily impregnated rim with a lighter centre. After centrifuging, the Golgi bodies were 'either unmoved or else tended to form a layer above the mitochondria, according to the length of time of centrifuging'. Golgi bodies of a similar type were described by Moore (1934) in *Blepharisma undulans*. Hall & Alvey (1933) described a typical 'vacuome' in *Colpidium colpoda*. Smyth (1941) was able to demonstrate the presence of an osmiophile (Golgi) cortex of the contractile vacuole of this species. Also in Weigl preparations a large mass of less densely staining osmiophile material was shown to be associated with the cortex and spread along the cell wall. During division this material became carefully separated into two blocks. An almost identical condition was also described by the same worker in *Plagiopyla* sp. In this case osmication was followed by examining the fixed ciliates at various times during the impregnation period. The Golgi cortex first showed up as a thin ring of small brown globules which later became black and swelled up to form an almost complete ring of black beads. On further impregnation the beads coalesced and the typical osmiophile cortex was obtained. In both *Plagiopyla* and *Colpidium* scattered osmiophile granules were present. In the same paper Smyth described osmiophilic cortices to both the contractile vacuoles of *Cyrtostomum leucas*, but the large associated mass was absent. In the recently discovered ciliate *Tillina canalifera* Turner (1940), using the standard methods, figured Golgi bodies present as spheres with granular osmiophile cortices and osmiophobe centres. The canals in this species were never impregnated. Park (1929) failed to impregnate the contractile vacuole of *Stentor coeruleus* and *Leucophrys patula*, and described the Golgi bodies as globules aggregated on the micronucleus.

Among parasitic ciliates most of the more common types have been investigated. Gatenby & King (1926) described densely osmiophilic bodies in the endoplasm of *Opalina ranarum*, and called them 'presumed Golgi bodies'. These were apparently attached to the base of the cilia—each cilium having an individual osmiophile body. Sokolska (1927) described disk-like elements in the endoplasm which he presumed represented Golgi bodies. These bodies had an outer lipoidiferous membrane forming a loop around an inner greyish substance. According to Sokolska, if the protozoan strikes an object the Golgi bodies in the region lose their disk-like shape and become irregularly rounded. Patten (1932), however, is opposed to the view that the structures described by Gatenby & King are Golgi bodies, as similar structures could be demonstrated by staining in iron alum haematoxylin after chrome-osmium fixation. She was unable to demonstrate any definite Golgi material by means of either of the Golgi-osmic techniques, and concluded that true Golgi material was probably absent in this ciliate. In the closely allied form *Protopalina* Richardson & Horning (1931) described vegetative granules and Golgi bodies as being present. The former were revealed by chrome-osmium preparations stained in iron alum haematoxylin; these granules they homologized with the osmiophile bodies of Gatenby & King. The Golgi bodies were revealed by Da Fano and Cajal silver methods as twisted rods and granules scattered in the cytoplasm.

MacLennan (1933) has given an account of the cytology of a number of *Oligotrichida* parasitic in cattle, namely, *Ophryoscolex*, *Epidinium*, *Ostracodinium*, *Polyplastron*, *Eudiplodinium*, and *Metadinium*. In all these forms Golgi granules are present in the cytoplasm and are described as accumulating around the contractile vacuole during diastole, later becoming reduced in number. Accessory vacuoles are formed in this granular region by solution of the granules in the vacuolar fluid. The whole mechanism is believed to represent an apparatus for the elimination of katabolic waste. The same worker (1936, 1937) described a somewhat similar condition in the fish parasite *Ichthyophthirius*. There were, however, two types of Golgi bodies in this form—endoplasmic Golgi bodies (intermediate lipoid bodies) and ectoplasmic Golgi bodies (excretory non-lipoidal granules). The latter were shown by both osmium and

silver techniques to be aggregated round the contractile vacuole during diastole in the parasitic stages. The endoplasmic Golgi bodies were few in the young parasite, but become more numerous as it grows, appearing later as hollow spheres, which MacLennan suggests take part in fat storage. No Golgi bodies of any kind were present during the encysted stages of the ciliate. In *Haptophyra*, an astomatous ciliate parasitic in the salamander *Hemidactylium scutatum*, Bush (1934) has described the Golgi material as a long osmiophile tube which is permanent and forms part of the vacuolar system which extends the full length of the ciliate. During division the tube is figured as dividing into two parts, each of which continues to function in the daughter cells. In *Nyctotherus* Horning (1927) described the Golgi bodies as clearly differentiated rod-like bodies or twisted filaments scattered throughout the cytoplasm and demonstrable by the Da Fano and Cajal techniques. Patten (1932) failed to demonstrate any inclusion identifiable as Golgi bodies and believed that structures described by Horning were bacteria, since they were stainable in Gram's stain.

IV. DISCUSSION

1. *Identification*

The majority of modern workers believe that the Golgi apparatus of the Protozoa is a definite cellular inclusion and that the osmic techniques of Weigl and Kolatchew in the hands of a careful worker will yield specific results (Gatenby, Duboscq, Grassé, King, Singh, Brown, Browne, Hill, Daniels, Smyth, *et al.*). Other cytologists have put forward some objections supported by a certain amount of evidence (Tennent, Gardiner, Smith, Hoerr, MacLennan, *et al.*). Kirkman & Severinghaus (1938 *a*, *b*, *c*), in their recent review on the metazoan Golgi apparatus, state: 'In spite of the imperfection of microchemical techniques, and notwithstanding an occasional dissenting voice, it seems likely that mistaking the identity of the Golgi apparatus in most vertebrate cells at least, is largely a thing of the past.' The position as yet is not so clear in the Protozoa, but agreement is gradually being reached as to the morphology and function of the Golgi material in the more common Protozoa. The 'vacuome' hypothesis of Parat and his associates, which for so long has confused the issue in the Protozoa, may now be considered obsolete in view of the very convincing work of Daniels (1938) on the ultracentrifuge. In general the Golgi apparatus described in Protozoa may be divided into three approximate groups for discussion: (*a*) those related to the base of the flagellum, i.e. the parabasal bodies; (*b*) those associated with the vacuolar system; and (*c*) those distributed throughout the cytoplasm in the form of granules, spheres, or dictyosomes.

2. *General morphology*

(*a*) *The parabasal body.* The hypothesis that the parabasal body of some flagellates represented the protozoan homologue of the metazoan Golgi apparatus was originally put forward by Gatenby & King (1923) and has since received considerable attention from Duboscq & Grassé (1925, etc.). The fact that the parabasal in many flagellates is osmiophilic and argentiophilic and secretory in some instances (*Jaenia*, *Tetramastix*, *Trichomonas*) seems to support this thesis. There are, however, difficulties in this homology. For example, in some instances impregnation takes place very slowly, even at 30° C. (60 days or more in *Trichomonas*, *Trypanosoma*, *Pyrsonympha*), in contrast with the usual ease of impregnation in metazoan cells. The parabasal is also stainable in Janus B green in high dilutions, and in this respect resembles mitochondria (*Tetramastix*, *Herpetomonas*, Grassé, 1926; *Crithidia*, Becker, 1923; *Trypanosoma*, Shipley, 1916). Brown (1930*a*) claimed to have demonstrated both parabasals and Golgi bodies* as separate entities in *Microjaenia* and is against the parabasal homology. Hill (1933) suggested that the bodies figured by Brown may be merely globules secreted by the parabasals. Parat (1928) also does not support the parabasal hypothesis. Hall (1930*a*) states 'there is still reasonable doubt as to the correctness of Grassé's view that the parabasal body is to be homologized with the Golgi apparatus of Metazoa'. The difficulty that the parabasal hypothesis cannot be extended to those flagellates which have no parabasal apparatus has been overcome to some extent by the suggestion of Gatenby (1938) that the parabasal in higher flagellates became associated with the contractile vacuole.

(*b*) *Golgi material associated with the vacuolar system.* The view of Nassonow (1924, 1925) that the osmiophile material *together with* the contractile vacuole represents the homologue of the Golgi apparatus in higher forms has led to much controversy. Nassonow based his homology on several criteria. First, the fact that the cortex to the contractile vacuole was osmiophilic after Golgi treatment. Secondly, there is very strong resemblance with the Golgi apparatus of certain cells of higher forms, especially sponge choanocytes and many germ cells. Thirdly, the osmiophile cortex is believed to bear the same relation to excretion as it does in higher cells. The great objection to Nassonow's homology is that Golgi bodies have been demonstrated in many forms lacking a contractile vacuole, e.g. the Sporozoa. On this account most modern workers do not accept Nassonow's hypothesis in its original form. Gatenby (1938) modified this thesis and suggested that the osmiophile material *alone* represented the Golgi apparatus and

* In *Vorticella*, however, Gatenby (1941) shows that apart from the osmiophile cortex of the contractile vacuole scattered comparable granules exist in the cytoplasm.

that its association with the contractile vacuole was only a secondary association. He further puts forward a view that links the parabasal type with that found in other forms. 'The Golgi apparatus developed in connection with the base of the flagellum in primitive flagellates, later becoming associated with the contractile vacuole, and when the necessity for an osmo-regulatory system of this type lapsed in the Metazoa, the osmiophile material persisted as an organ associated in some way with the removal of water from secretion products which in many cells arise in close propinquity to the Golgi apparatus.' This attractive hypothesis explains to some extent the diverse kinds of Golgi material found in the various classes of Protozoa. With regard to the osmiophile cortex described by many authors there is growing evidence that the structure of this is not as simple as many writers would believe. MacLennan (1941) points out that most workers use the warm method of impregnation advocated by Nassonow, and that this technique is liable to produce over-impregnation. In the Ophryoscolecidae, MacLennan (1933) found that the warm method produced a thick osmiophile membrane around the contractile vacuole, whereas the cold method demonstrated this same region to be granular. This writer has also pointed out that the solid cortices of the vacuolar walls figured in *Chilodon* and *Dogielella* (Nassonow, 1924, 1925) show a definite granular roughening of the outer region, and suggests that the solid membrane described is merely the result of the over-impregnation of a granular zone. Smyth (1941) showed that in *Plagiopyla* the osmication of the cortex results first in the appearance of a thin ring of brown granules which later become black and swell up to form an almost complete circle of black globules; further impregnation results in the formation of the typical osmiophile cortex so often described. Willis, working on *Lagenophrys* (quoted by Gatenby, 1941), also describes, after osmication, 'blackening of the accumulated perivacuolar granules which give the appearance of a distinctive Golgi cortex to the contractile vacuole'. MacLennan (1941), reviewing the Golgi material in *Epidinium*, *Euplodinium*, *Ichthyophthirius* and *Metadinium*, writes that 'impregnation of parts of the vacuolar system is due to the aggregation of osmiophilic granules around the fluid vacuoles and their membranes'. This writer believes that the contractile vacuole membrane only impregnates 'in extremely specialized cases' and quotes *Paramecium* (Nassonow, 1924; Von Gelei, 1928) and *Haptophyra* (Bush, 1934) as examples. Gatenby (1941) has shown that in *Vorticella* the original osmiophile cortex passes over completely to one daughter cell during division and that the new one is formed by migration of the scattered Golgi bodies. Von Gelei (1933) observed in *Spathidium* that in addition to the osmiophile wall of the contractile vacuole, smaller osmiophile vacuoles are found in its vicinity and scattered in the cell. Weatherby (1941) suggests that this may represent 'the origin of contractile vacuoles by the coalescence of secondary vacuoles which have arisen in more or less remote parts of the organism'.

(*c*) *Scattered Golgi elements.* In view of the apparent granular nature of the osmiophile cortex of many organisms it seems likely that this third type of Golgi body is not as distinct as was formerly believed. The fact that it so frequently occurs in organisms which also have a vacuolar cortex (e.g. *Colpidium*, *Plagiopyla*, *Chilomonas*, *Syctomonas*, *Vorticella*, etc.) indicates how closely related the two types are. In the Sporozoa the Golgi bodies frequently exhibit an osmiophile and osmiophobe region (Hirschler, 1914; Gatenby & King, 1923; Daniels, 1938; Subramaniam & Ganapti, 1938). A similar condition has been described in many ciliates (Gatenby & King, 1923; Moore, 1934; Browne, 1938; Turner, 1940; Smyth, 1941). In most other organisms described, the diffuse type of Golgi body consists simply of uniformly staining osmiophilic spheres. Subramaniam & Ganapti (1938) believe that the dictyosome structure is formed from the granular Golgi body. They write: 'A Golgi granule when it enlarges becomes differentiated into a vesicle having chromophile and chromophobe regions. Rupture of the vesicle gives rise to batonettes in which the chromophobe part is in relation with the cytoplasm.' MacLennan (1936, 1937) believes that the granular type is divisible into endoplasmic Golgi bodies (intermediate lipoid bodies) and ectoplasmic Golgi bodies, but this statement has not so far been confirmed by other workers. Some workers (Kraschenninikow, 1929; MacLennan, 1933) have claimed that the Golgi granules accumulate around the contractile vacuole only during diastole in *Epidinium* and *Eudiplodinium*; a similar arrangement has been described in *Ichthyophthirius* in the parasitic stages, but in this organism the osmiophile granules are not present during encystment. Von Gelei (1938) has shown that in some ciliates the granules are continually migrating towards the vacuoles, but never away from them.

In the Rhizopoda the granular complex is confusing. The claims of Brown (1930*b*) and Mast & Doyle (1935 *a*, *b*) to have demonstrated Golgi bodies in *Amoeba* have not been confirmed by Singh (1938) who supplemented his work with the application of the ultra-centrifuge. The results of Hall & Loefer (1930) on *Euglypha* have never been challenged; but the claim that the Golgi granules are argentiophilic is disputable on the grounds that the silver technique has met with such universal failure in all other classes of Protozoa. Gatenby (1941), from results of work on *Arcella* and *Nebela*, believed that the Golgi apparatus is absent in these forms. A similar condition is possibly found in *Stylonychia* (Smyth, 1941) and *Nyctotherus* (Patten, 1932), among ciliates.

3. *Function*

In view of the great diversity of structure found in the types of Golgi bodies in the Protozoa, any hypo-

thesis regarding function is difficult. In general the Golgi bodies are concerned with secretion, according to most authors. Duboscq & Grassé (1925, etc.) have shown the secretion of globules from the parabasals of many flagellates. Nassonow (1924–5) believed that the osmiophile cortex bears the same relation to excretion as does the metazoan Golgi apparatus to excretion. He suggested that the cortex accumulates water and other waste materials which are shed into the contractile vacuole. He believed the process takes place in two phases: first, a 'bound secretion phase', at which the water and the excretory substances dissolved in it form as vacuoles in the interior of the osmiophile cortex; and secondly, a 'free secretion phase' during which the small vacuoles flow together to form a true contractile vacuole. MacLennan (1933) believes a similar process takes place in the Ophryoscolecidae. He writes: 'The vacuolar region found in these ciliates shows definite evidence of the elimination of materials by means of the vacuolar fluid and corresponds to the secretory region or "region of Golgi" in gland cells.' This same writer (1936) showed that the vacuoles are osmiophilic only during the active feeding stages and not during encystment (when active feeding ceases), and claimed this to indicate that the osmiophile (Golgi) material is concerned with excretion of metabolic waste. Gatenby (1938) points out that if we consider that an association between primitive Golgi material (i.e. parabasal) and contractile vacuole has occurred, an excretory function would be the only one that would be of advantage to the protozoan. Although an excretory function is agreed on in many forms by most cytologists, the exact mechanism of this process is doubtful. Apart from the views of Nassonow and MacLennan mentioned above, Smyth (1941) points out that one of the properties of lipoid substances is that they have the power of forming a film on liquid surfaces; this may act as a membrane, the permeability of which is easily affected by the presence of various substances. Applying this theory to the osmiophile cortex he suggests the possibility that it acts as a sieve or filter which prevents the escape of certain necessary substances and at the same time allows the passage of unwanted water. There is much evidence to indicate that the contractile vacuole is an organ related to the control of osmotic pressure (Zolger, Wolff, Yocom, Kitching, *et al.*), and it seems possible that the supposed excretion of water by the Golgi material also plays a part in this regulation of osmotic balance. The fact that the osmiophile cortex is absent in many of the marine Protozoa (*Euplotes*, Smyth, 1941; *Oxyrrhis*, Gatenby, 1941) seems to support this view. On the other hand, if this hypothesis were correct we should not expect to find osmiophile material associated with the vacuolar system in parasitic Protozoa, owing to the high osmotic pressure of the fluid in the internal body cavities. Yet a very definite osmiophile complex has been described in *Haptophyra* (Bush, 1934), *Ichthyophthirius* (MacLennan, 1936–7), and *Copromonas* (Gatenby & Singh, 1938), to mention only a few instances.

The osmoregulatory theory also fails to explain the absence of Golgi material associated with the contractile vacuole in many fresh-water Protozoa, as *Stentor*, *Leucophrys* (Park, 1929), *Spirostomum* (Browne, 1938), *Blepharisma* (Moore, 1934), *Lionotus* (Smyth, 1941). It is even more difficult to account for the *complete* absence of any kind of Golgi substance in other fresh-water forms such as *Nebela* (quoted by Gatenby, 1938), *Amoeba* (Singh, 1938), *Stylonychia* (Smyth, 1941). It is interesting, however, to note that in some forms possessing osmiophile material associated with the vacuole, when division takes place the Golgi substance passes over whole to *one* daughter organism in a small percentage of cases, e.g. in *Copromonas* (Gatenby & Singh, 1938), *Chilomonas* (Gatenby & Smyth, 1940), *Rhabdomonas* (Smyth, 1944). In *Vorticella* (Gatenby, 1941) this occurs in every instance during normal division: from which we may conclude that the Golgi material associated with the contractile vacuole has not such an important role to play in the general metabolism of the cell, since the cell is able to live for some time without it.

In the Sporozoa, in which the Golgi bodies are usually in the form of scattered granules or dictyosomes, Joyet-Lavergne (1926) has suggested that they take some part in the formation of fatty substances. MacLennan (1936) suggests a similar function for the Golgi bodies in *Ichthyophthirius*.

V. SUMMARY

(1) The Golgi apparatus is considered to be a definite cellular inclusion in the Protozoa, although there is still considerable disagreement regarding identification, morphology, distribution, and function in the different classes. (2) The osmic techniques of Weigl (Mann-Kopsch) and Kolatchew, followed by bleaching in hydrogen peroxide or turpentine, are specific for the Golgi material in most organisms. The silver methods of Cajal and Da Fano are seldom successful in the Protozoa. The application of the ultra-centrifuge has proved that the neutral-red bodies, i.e. the 'vacuome', are not identical with the Golgi bodies. (3) In the Mastigophora the parabasal bodies of many forms are considered to represent the Golgi apparatus, though the homology may not be said yet to be fully proved. In other flagellates the cortex of the contractile vacuole is impregnated by the osmic Golgi techniques and its behaviour during division in many organisms is similar to that of the metazoan apparatus. These criteria seem sufficient to identify it as Golgi material. (4) In the Sporozoa there is general agreement that the Golgi apparatus is represented by scattered globules or dictyosomes, often possessing osmiophile and osmiophobe regions. In some cells it shows the juxtanuclear position so characteristic of the Golgi apparatus of many metazoan cells. The homology is also supported by the fact that in the centrifuged sporozoan the osmiophile material occupies the same position relative to the

other inclusions as does the Golgi material in higher animal cells. (5) There is still much confusion as to the exact nature of the Golgi apparatus in the Rhizopoda, and there is little agreement as to its identity in any of the types studied. Some workers have described the Golgi apparatus as being represented by scattered osmiophilic globules or granules. Others have failed to identify any Golgi material in members of this class. (6) In the Ciliata most workers agree that the osmiophile cortex to the contractile vacuole found in many forms represents the Golgi apparatus. In some instances scattered Golgi bodies are distributed throughout the cytoplasm. These are often present in combination with the Golgi cortex. There is no convincing evidence that any Golgi material is present in sea-water ciliates, the fresh-water form *Stylonychia*, or the parasitic *Nyctotherus*. (7) Nassonow's view that the osmiophile material *together with* the contractile vacuole represented the Golgi apparatus is not supported by modern workers. The osmiophile material *alone* is believed to represent this inclusion. There is much evidence to suggest that there is a single type of granular Golgi body, from which both the osmiophile cortex type and the dictyosome type are formed. It is likely that the Golgi apparatus arose in connexion with the base of the flagellum, later becoming associated with the vacuolar system in many forms. (8) In view of the great diversity in form of the protozoan Golgi apparatus any single hypothesis regarding function is inadequate to explain all the facts. There is much evidence, however, to indicate that it is concerned with the mechanism of excretion or secretion, and possibly also with the osmoregulation of the organism. In the Sporozoa it may take part in fat metabolism.

I wish to thank Prof. J. Brontë Gatenby of the Zoological Department, Trinity College, Dublin, and Dr Geoffrey Bourne of the Physiological Department, Oxford University, for reading through the manuscript.

VI. REFERENCES

BAKER, C. L. (1933): *Arch. Protistenk.* **80**, 434.

BEAMS, J. W. (1930): *Rev. Sci. Instrum.* **1**, 667.

BEAMS, J. W., WEED, A. W. & PICKLES, E. G. (1933): *Science*, **78**, 338.

BECKER, E. R. (1923): *J. Parasitol.* **9**, 141.

BOURNE, G. (1942): *Cytology and Cell Physiology*, chap. IV. Oxford Univ. Press.

BROWN, V. E. (1930*a*): *Univ. Calif. Publ. Zool.* **36**, 67. — (1930*b*): *Biol. Bull. Woods Hole*, **59**, 240.

BROWNE, K. (1938): *J. R. Micr. Soc.* **58**, 188.

BUSH, M. (1934): *Univ. Calif. Publ. Zool.* **39**, 251.

CAUSEY, D. (1925): *Univ. Calif. Publ. Zool.* **28**, 1

CHATTON, E. & GRASSÉ, P. (1929): *C.R. Soc. Biol., Paris*, **100**, 281.

COWDRY, E. V. & SCOTT, G. H. (1928): *Arch. Inst. Pasteur Tunis*, **17**, 233.

DANIELS, M. (1938): *Quart. J. Micr. Sci.* **80**, 293.

DUBOSCQ, O. & GRASSÉ, P. (1925*a*): *C.R. Soc. Biol., Paris*, **92**, 154. — (1925*b*): **93**, 345. — (1927): **96**, 92. — (1928): **99**, 1118. — (1933): *Arch. Zool. exp. gén.* **73**, 381.

DUNIHUE, F. W. (1931): *Arch. Protistenk.* **75**, 476.

FAURÉ-FREMET, E. (1925): *C.R. Soc. Biol., Paris*, **93**, 500.

GATENBY, J. B. (1930): *J. R. Micr. Soc.* **50**, 20. — (1931): *Amer. J. Anat.* **48**, 421. — (1938): *Evolution of the Cell.* In *Essays addressed to Prof. Goodrich.* Oxford Univ. Press. — (1941): *Proc. Roy. Irish Acad.* B, **46**, 161.

GATENBY, J. B. & KING, S. D. (1923): *Quart. J. Micr. Sci.* **67**, 387. — (1926): **70**, 217.

GATENBY, J. B. & SINGH, B. N. (1938): *Quart. J. Micr. Sci.* **80**, 567.

GATENBY, J. B., SINGH, B. N. & BROWNE, K. (1938): *La Cellule*, **47**, 229.

GATENBY, J. B. & SMYTH, J. D. (1940): *Quart. J. Micr. Sci.* **81**, 595.

GELEI, J. VON (1928): *Arch. Protistenk.* **64**, 446. — (1933): **81**, 231. — (1938): **90**, 368.

GRASSÉ, P. P. (1925): *C.R. Acad. Sci., Paris*, **181**, 482. — (1926): *Arch. Zool. exp. gén.* **65**, 345.

HALL, R. P. (1929): *J. Morph.* **48**, 105. — (1930*a*): *Arch. Protistenk.* **69**, 7. — (1930*b*): *J. Morph.* **49**, 139. — (1931*a*): *Ann. Protistol.* **69**, 7. — (1931*b*): *Z. Zellforsch.* **13**, 770.

HALL, R. P. & ALVEY, C. H. (1933): *Trans. Amer. Micr. Soc.* **52**, 1.

HALL, R. P. & DUNIHUE, F. W. (1931): *Trans. Amer. Micr. Soc.* **50**, 196.

HALL, R. P. & LOEFER, J. B. (1930): *Arch. Protistenk.* **72**, 365.

HALL, R. P. & NIGRELLI, R. F. (1930): *Trans. Amer. Micr. Soc.* **49**, 18.

HAYE, A. (1930): *Arch. Protistenk.* **70**, 1.

HILL, J. C. (1933): *J. R. Micr. Soc.* **53**, 227.

HIRSCH, G. C. (1939): *Form und Stoffwechsel der Golgi-Körper.* Protoplasma Monographs, Berlin.

HIRSCHLER, J. (1914): *Anat. Anz.* **47**, 289. — (1924): *C.R. Soc. Biol., Paris*, **90**, 891. — (1927): *Z. Zellforsch.* **5**, 704. — (1932): *C.R. Acad. Sci., Paris*, **194**, 1682.

HOERR, N. L. (1936): *Anat. Rec.* **66**, 149.

HORNING, E. S. (1927): *Austr. J. Exp. Biol. Med. Sci.* **4**, 69.

HYMAN, L. (1940): *Protozoa through Ctenophora.* McGraw-Hill.

JOYET-LAVERGNE, P. (1923): *C.R. Acad. Sci., Paris*, **177**, 975. — (1924): **178**, 128. — (1925): *Arch. Anat. micr.* **22**, 1. — (1926): *C.R. Soc. Biol., Paris*, **94**, 930.

KING, S. D. (1926): *Quart. J. Micr. Sci.* **70**, 147. — (1927): *J. R. Micr. Soc.* **48**, 342.

KING, R. (1933): *Trans. Amer. Micr. Soc.* **52**, 103. — (1935): *J. Morph.* **58**, 555.

KIRKMAN, H. & SEVERINGHAUS, A. E. (1938*a*): *Anat. Rec.* **70**, 413. — (1938*b*): **70**, 557. — (1938*c*): **71**, 79.

KRASCHENINNIKOW, S. (1929): *Z. Zellforsch.* **8**, 470.

LWOFF, M. & LWOFF, A. (1929): *C.R. Soc. Biol., Paris*, **100**, 557.

MACBRIDE & HEWER (1931): In *Recent Advances in Cytology.* London: Churchill.

MACLENNAN, R. F. (1933): *Univ. Calif. Publ. Zool.* **39**, 205. — (1936): *Arch. Protistenk.* **86**, 403. — (1937): *J. Exp. Zool.* **76**, 423. — (1940): *Trans. Amer. Micr. Soc.* **59**, 149. — (1941): In *Protozoa in Biological Research.* Columbia Univ. Press.

MAST, S. O. & DOYLE, W. L. (1935*a*): *Arch. Protistenk.* **86**, 155. — (1935*b*): **86**, 278.

MOORE, I. (1934): *J. Exp. Zool.* **69**, 59.

NASSONOW, D. (1924): *Arch. mikr. Anat.* **103**, 437. — (1925): *Z. Zellforsch.* **2**, 87.
PARAT, M. (1928): *Arch. Anat. Micr.* **2**, 87.
PARAT, M. & PAINLEVÉ (1924): *C.R. Acad. Sci., Paris*, **179**, 612.
PARK, O. (1929): *Trans. Amer. Micr. Soc.* **48**, 20.
PATTEN, R. (1932): *Proc. Roy. Irish Acad.* B, **41**, 73.
PATTEN, R. & BEAMS, H. W. (1936): *Quart. J. Micr. Sci.* **78**, 615.
RICHARDSON, K. C. & HORNING, E. S. (1931): *J. Morph.* **52**, 27.
SAEDELEER, H. (1930): *C.R. Soc. Biol., Paris*, **103**, 160.
SHIPLEY, P. G. (1916): *Anat. Rec.* **10**, 439.
SIGOT, A. (1931): *C.R. Soc. Biol., Paris*, **106**, 1069.
SINGH, B. N. (1938): *Quart. J. Micr. Sci.* **80**, 601.
SMYTH, J. D. (1941): *Proc. Roy. Irish Acad.* B, **46**, 189. — (1943): *Nature, Lond.*, **151**, 110. — (1944): *Quart. J. Micr. Sci.* (in the Press).
SOKOLSKA, J. (1927): *C.R. Soc. Biol., Paris*, **96**, 570.
SUBRAMANIAM, M. K. & GANAPATI, P. N. (1938): *Cytologia, Tokyo*, **9**, 1.
TENNENT, D. H., GARDINER, M. S. & SMITH, D. E. (1931): *Pap. Tortugas Lab.* **27**, 1.
TURNER, J. P. (1940): *Arch. Protistenk.* **93**, 255.
TUZET, O. (1931): *Arch. Zool. exp. gén.* **71**, 16.
WALKER, C. (1928): *Proc. Roy. Soc.* B, **103**, 397.
WALKER, C. & ALLEN, M. (1927): *Proc. Roy. Soc.* **101**, 468.
WEATHERBY, J. H. (1941): *Protozoa in Biological Research.* Columbia Univ. Press.

NUCLEAR AND SOMATIC PHASES IN THE FLORIDEAE

BY KATHLEEN M. DREW, D.Sc.
Department of Cryptogamic Botany, University of Manchester

(*Received* 2 *May* 1944)

A perusal of recent literature not only affords striking evidence of considerable diversity in the nuclear and somatic phases of the red algae but directs attention to the superficiality and inadequacy of our present knowledge. In some orders of the Florideae, for example, the field is practically untouched. Indeed, one cannot fail to note how few in all have been the workers, and still fewer the number who have devoted many consecutive years of work to a group, which is as interesting as it is varied and extensive. Moreover, the assumption that all floridean life histories fit into one or other of only two schemes is no longer tenable, an assertion in support of which both morphological and cytological evidence can be adduced. While the primary need is thus seen to be for a considerable extension of critical morphological and cytological investigation, there is, in my view, scope also for the re-examination of material, collected over a wide geographical range, of those species which have come to be regarded as affording classical examples of life histories within the group. Experimental work, particularly on the factors governing growth and reproduction, is also desirable. While stressing the need for cytological evidence it is realized that such information for any one alga represents only a part of the picture. But it must also be pointed out that, in the absence of the corresponding cytological data, the interpretation of morphologically established facts can be misleading; in some instances interpretations based on them alone have eventually proved to be untenable.

The recent extension of our knowledge of the Florideae makes some kind of stocktaking desirable in the interest of the further advancement of the subject. Accordingly an attempt has been made to prepare a critical connected statement of the established cytological facts, together with other observations, the effect of which is to show the need of further cytological investigation. This account will be given in a systematic rather than a chronological order, the classification followed being that of Kylin (1937).

I. TERMINOLOGY

In the literature of the Florideae, and indeed in that of the algae in general, many terms such as generation, alternation of generations, phase, life cycle and so on are of frequent occurrence. Their precise meaning cannot be inferred from the vague way in which they are used and may vary from author to author, as it has done so obviously in the case of 'alternation of generations'. Moreover, some of the terms have been brought into common use without having been defined in the first instance. The additional value gained by a clear definition of terms at the outset cannot be too highly assessed.

The terminology used in other plant groups, which were investigated earlier than the algae, should not be accepted on *a priori* grounds as being suitable for the algae. The incautious use of these terms inevitably tends to introduce an element of prejudgement regarding the interpretation of the new data. For example, the term 'generation', a word of various meanings (*A New English Dictionary*, 1901), has little or no meaning in the algae, whether it is used in the sense of the total of those somatic

phases covered by a nuclear cycle or in the restricted sense of the equivalent of the gametophyte or sporophyte of a fern, for instance. It would seem better to use the word *phase*—having in mind that one succeeds another—for both the cytological and morphological periods distinguishable in the life history of a given alga. As used in this article it has the meaning given in the dictionary mentioned (1909): 'Any one aspect of a thing of varying aspects; a state or stage of change or development.' I suggest that for the present *life history* is the most satisfactory of the terms in use for expressing the sum total of our knowledge of the sequence of phases shown by any one alga. A life history can be considered in terms of its *somatic* and *nuclear phases*;* indeed, these phases may have the same limits sometimes, but this is not invariable. Thus in its comprehensive aspect a life history may be considered to consist of x somatic phases and y nuclear phases; in many instances x and y are of the same numerical value, but that this is not always so is indicated by the Florideae. Certain somatic phases are such that they may be repeated once or several times before the change to the succeeding phase occurs, as in some members of the Ectocarpales. Thus, if *A* represents one type of somatic phase and *B* another type, the sequence may be *AB* or *AAB* or *AAAB* or *ABB*, and so on. In the Florideae this repetition of a phase is not very common but does occur in those species possessing monospores as well as sexual reproduction. There are species the life history of which appears to consist of the repetition of a single phase, e.g. *Rhodochorton* spp. which reproduce by means of monospores only. Life history is therefore definable as *the observed sequence or sequences of somatic and nuclear phases known for the species under consideration.* In nature a life history may or may not be cyclic† and possibly may vary from locality to locality or season to season. Thus if the sequence of phases is *AAB* in one locality or season it may be *ABB* in another. Further investigations may show these concepts to be inadequate: meanwhile, this appears to be the most objective method of describing the facts at present available.

The nuclear phases are usually cyclic, and there is in the majority of species investigated so far a regular alternation between the *haplo-* and *diplophase*, but important exceptions are known. The limits of the nuclear phases can be ascertained by counting the number of chromosomes present at the various nuclear divisions, and it is highly essential to obtain this information for as many species as possible. In addition, the nuclear phases are less easily influenced by external factors than are the somatic phases.

Because of the occurrence of the cystocarp the definition of somatic phases in the Florideae presents difficulties. The sexual plant, the *gametophyte*‡ (with or without asexual spores), clearly represents one phase, and the tetrasporic (i.e. tetraspore-bearing) plant, the *tetrasporophyte*,‡ of the *Polysiphonia* type another, but the *cystocarp* or *carposporophyte* has been variously interpreted. This fact occasions no surprise since the post-fertilization developments are so varied in this group. In contrast to Svedelius (1927) who thinks it better not to term the floridean cystocarp a generation, Kylin (1937) considers it a generation from the morphological standpoint. Oltmanns (1898, 1904 and 1923) called the cystocarp a sporophyte and the tetrasporophyte a 'Nebenfruchtform', but this view has not been generally accepted. Kniep (1928), following Janet (1914), terms it a carposporophyte and Smith (1938) considers it a phase. In spite of the non-homogeneity of the floridean cystocarp, it originates from and ends with the single-celled state, and this would seem to be a good criterion for the limits of a phase.§ So, although no parallel occurs in other plant groups, I consider the cystocarp as a distinct and separate phase peculiar to this group. There is, in fact, a greater morphological difference between the gametophyte and the cystocarp than between the gametophyte and the tetrasporophyte.‖ This is particularly true in species such as *Spermothamnion Turneri*, where the reproductive organs which characterize these two phases exist side by side on the same plant, or even branchlet. On the other hand, in *Plumaria elegans* Drew (1939) a fourth *somatic phase* (the triploid nuclear phase) is known, this being characterized by the formation of *paraspores*.

It was Svedelius (1915) who first pointed out that the succession of somatic phases in the Florideae was different from that of other plant groups and that the somatic and nuclear phases were not coincident. He gave the name *haplobiontic* to the type of life history exemplified by *Scinaia furcellata* and *diplobiontic* to that of *Polysiphonia violacea*. Svedelius defined these terms and in 1931 amplified the definitions and condoned the use of the terms for other groups of plants. It is to be inferred from these writings of Svedelius that the unit or bionte is a physiological unit irrespective of its morphology

* Kylin (1916*b*, 1917*a*) clearly distinguishes between somatic and nuclear phases. He retains the word generation for the former and phase for the latter. Other writers have used this terminology.

† As the life history is not always cyclic, *life cycle* is not a suitable term for general use. The German writers use the word *Wechsel* (=change, exchange), e.g. Biontonwechsel Kernphasenwechsel, Generationswechsel; this is often much more appropriate than cycle.

‡ These terms seem suitable and are self-explanatory.

§ This would not hold for 'juvenile' forms, such as the 'Chantransia' stage of *Batrachospermum*, if considered a somatic phase. This is not the place, however, for the discussion of that point.

‖ Structural dimorphism between the tetrasporophyte and gametophyte is known in the genus *Galaxaura*.

or cytology, e.g. both *Scinaia furcellata* and the Bryophyta are haplobionts. In an earlier paper (Drew, 1943) I suggested 'that the terms were kept for the Florideae and based on cytological criteria by which *Plumaria elegans* would be a triplobiont and *Spermothamnion Turneri* a polybiont'. After further consideration it seems undesirable to extend the use of the terms haplobiontic and diplobiontic or to redefine them for use in a scheme of classification of life histories based on somatic and nuclear phases. For example, should *Liagora tetrasporifera* be shown to have reduction division in the tetracarposporangium, Svedelius would still class it as haplobiontic which would not differentiate it from a type such as *Scinaia furcellata*. It is to be noted that Børgesen (1927) would consider such a life history as diplobiontic. It should be remembered that the definitions were made by Svedelius at a time when our knowledge of the Florideae was much more incomplete than at present.

Smith (1938), in discussing floridean life histories, uses another terminology; thus 'the following three types of life cycle may be recognized in the Florideae: a biphasic alternation of a gametophyte and a haploid carposporophyte; a biphasic alternation of a gametophyte and a diploid carposporophyte; and a triphasic cycle in which both carposporophyte and tetrasporophyte are diploid'. Although there is no cytological proof extant of the second type, and other types of life histories exist, this statement is adequate from the morphological point of view. Only by implication, however, does it convey the information that the nuclear cycle is diphasic. No single word can adequately describe the somatic and nuclear phases of any one life history, and therefore it would seem more exact to state that *the life history is morphologically mono-, di-, tri-, or polyphasic and cytologically mono-, di-, tri-, or polyphasic.* For example, *Scinaia furcellata* is morphologically and cytologically diphasic, *Polysiphonia violacea* morphologically triphasic and cytologically diphasic, *Spermothamnion Turneri* cytologically polyphasic and *Plumaria elegans* morphologically tetraphasic and cytologically triphasic. We have sufficient evidence already to know that, from the morphological point of view, there is more than one type of diphasic life history, and it is possible there are various types of diphasic nuclear cycles also. Just as some of the monophasic types may be primitive while others exemplify a derived condition, so species with two somatic phases may include some primitive and some derived forms. When this derived condition has been established such species should obviously be separated from the others, and this could be done by designating them *derived monophasic* or *derived diphasic* in contrast to those initially mono- or diphasic.

The preceding remarks will have served incidentally to emphasize the difference between the life histories of the algae (the Florideae in particular) and those of the Pteridophyta (excepting instances of apogamy and apospory) and Bryophyta, where there is a regular, obligatory, cyclic alternation of two 'generations', morphologically very distinct, and where the nuclear phases and somatic phases are coincident and well defined.

II. SURVEY OF THE PRESENT POSITION

1. Order I. Nemalionales

This order, of eight well-defined families comprising about 250 species, contains forms considered to be the most primitive of all the Florideae. The cystocarp varies from an exceedingly simple structure in some families to one of considerable complexity in others. One marked characteristic of the order is, generally speaking, the absence of tetrasporangia. While this is true for the majority of the species, tetrasporangia are known for either certain or all species of various genera belonging to four out of the eight families. They develop either in the cystocarp or on separate plants. No such species have been investigated cytologically, and so it is still unknown whether these sporangia are associated with any change in the chromosome number. In other genera monosporangia occur, either on the juvenile or the adult forms, and probably reproduce the gametophyte or tetrasporophyte directly. Howe (1920) describes the formation of monosporangia-bearing disks, arising from the terminal or subterminal cells, in the genus *Liagora*.

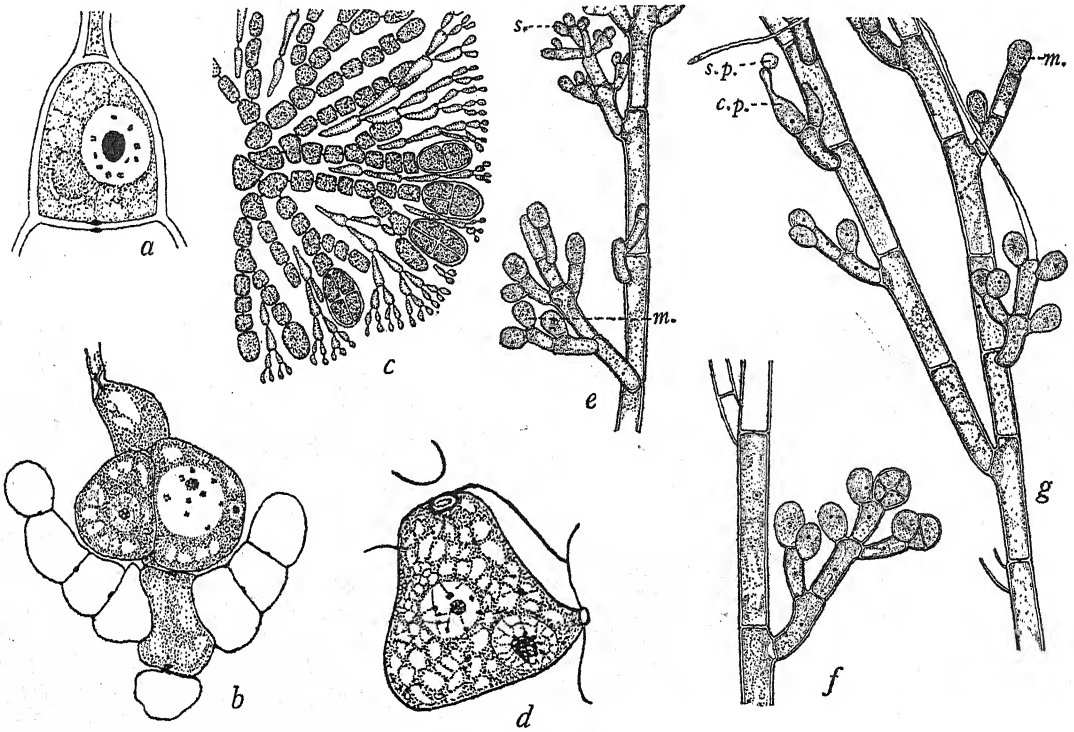
Our present morphological knowledge indicates the existence of a great variety of types of life history in this order. Examples of such with one, two or three somatic phases, sometimes more than one type occurring within the limits of a single family or genus, are all known. Cytological information is very scanty and confined to three examples of one type of diphasic life history, where the diploid nuclear phase is of such short duration that there is no corresponding morphological phase, both of the somatic phases being haploid. Information available suggests that cytological investigations of other types, particularly species with tetrasporangia, would provide results having considerable theoretical importance. Using *Scinaia furcellata* Svedelius (1915) was the first to demonstrate the occurrence of reduction division immediately after fertilization in the Florideae. This was followed by Kylin's work on *Nemalion multifidum* (1916) and *Batrachospermum moniliforme* (1917) and that of Cleland on *Nemalion multifidum* (1919).

In *Rhodochorton* (Chantransiaceae), much-needed cytological investigations are hampered by the small size of the nuclei. The majority of the species reproduce by means of monospores only, but at the other extreme there are two species, one marine (*R. efflorescens*) and another fresh water (*R. violaceum*) (Fig. 1 *e, f, g*), for which plants bearing sexual organs in addition to others bearing tetrasporangia are known

Drew (1935). Monosporangia occur plentifully as well, but on special plants in the case of *R. efflorescens*. Between these two extremes there are various intermediate types, such as species with sexual reproduction and monosporangia,* and others with tetrasporangia (with or without monosporangia) but no sexual reproduction. Bisporangia or polysporangia characterize other species, and it is apparent, therefore, that more than one type of life history, from a morphological point of view, is represented in this one genus (cf. Drew, 1928, p. 159), but whether the nuclear phases are equally varied is unknown. The monotypic fresh-water genera, of uncertain systematic position, *Thorea* and *Nemalionopsis* (Skuja, 1934) reproduce by monospores only as far as is known at present and are therefore monophasic.

Kylin (1917) figured the prophase stage of the first division of the fusion nucleus of *Batrachospermum moniliforme* (Batrachospermaceae), which he interpreted no doubt correctly as diakinesis. The number of gemini appears to be 10, the number also seen in the mature spermatium. The life history in this species is therefore of the *Scinaia* type. In this genus fertilization has been observed by several workers, the fusion of the male nucleus from the spermatium with the female nucleus of the carpogonium having been clearly figured by Osterhout (1900) for *B. Boryanum*. As the sequence of the first cell and nuclear divisions in the carpogonium is unknown it is not clear whether a tetrad of nuclei results from reduction division. However, it seems likely that all daughter nuclei take part in the formation of the cystocarp.

No cytological work has been done on the neighbouring family of fresh-water algae, the Lemaneaceae, but species of the Helminthocladiaceae have received some attention. Through the work of Kylin (1916) it is known that the life history of *Nemalion multifidum*† resembles that of *Scinaia furcellata*. He presented clear evidence of a diakinesis stage (with 10 gemini) (Fig. 1*a*) in the first division of the nucleus of the fertilized carpogonium. Additional evidence that this is the place of reduction division in the life history is provided by the fact that 8–10



Text-fig. 1. Diakinesis in carpogonium of *a*, *Nemalion multifidum* (× 1770, Kylin) and hypogynous cell of *b*, *Scinaia furcellata* (× 1440, Svedelius); and *d*, *Asparagopsis aramata* (× 1700, Svedelius); *c*, tetracarposporangia of *Liagora tetrasporifera* (Kylin). Monosporangia on *e*, male; *g*, female; and *f*, tetrasporic plants of *Rhodochorton violaceum* (× 500, Drew).

* This is the condition in *Balbiania investiens* (Sirodot, 1876).

† This alga had been investigated previously by Wolfe (1904) who found approximately 16 chromosomes present in the cells of the cystocarp up to the period of spore formation and approximately 8 in the cells of the thallus, 'reduction division being immediately associated with the production of the carpospores'.

chromosomes were observed in the mature spermatium and approximately 10 in the cells of the gonimoblast. No vegetative divisions were seen however. Only one of the daughter nuclei of the first division is used in the development of the sporogenous filaments. A longer account of the cytology of this same alga is given by Cleland (1919), who confirmed Kylin's findings in all essentials, saw fertilization clearly and figured the prophase stages of the first nuclear division following it. Cleland considered this a meiotic division, and although the figures given may be variously interpreted, the fact that the haploid number occurs in all other divisions seen, supports his conclusion. Two species, *Liagora tetrasporifera* and *Helminthocladia Hudsoni*, belonging to the same family, the Helminthocladiaceae, differ from *Nemalion multifidum* and other species of the genera to which they belong in that the protoplast of the carpogonium divides into four and thus resembles a tetrasporangium (Fig. 1*c*). There is no information regarding the nuclear phases, and without this we can only speculate (and possibly incorrectly) regarding the life history. The development of the carpogonial branch and gonimoblast in these species show no unusual features. While the division of the carposporangium is constant in *Liagora tetrasporifera* and *Helminthocladia Hudsoni* it is known as a rare occurrence in other species. I have seen once-divided carposporangia* while investigating a species of *Batrachospermum*, belonging to the *Helminthosum* section of the genus, and Westbrook (1930*b*) states: 'In one female plant' (of *Composthamnion thuyoides*, Ceramiaceae) 'which was sectioned, two tetrasporangia of normal appearance were found among a mass of carpospores.'

In the Chaetangiaceae, *Scinaia furcellata* and the genus *Galaxaura* call for comment. In *Scinaia furcellata* the first division of the fusion nucleus was found by Svedelius (1915) to be a reducing one, as the figures of diakinesis (Fig. 1*b*) show to be undoubtedly the case. The fusion nucleus (fertilization was not observed) moves into the hypogynous cell before dividing, and of the four nuclei which result one initiates the development of the gonimoblast through the carpogonium. These observations form a part of the author's detailed investigations, only a portion of which are cytological. No count of the number of chromosomes in the somatic nuclei was obtained and in the majority of prophases figured, excepting diakinesis, chromatin threads still connect the bodies considered to be chromosomes. The haploid chromosome number is given as 10, based on the count at diakinesis, but in the spermatia, carpogonial branch and carposporangium, 7, 8 and 9 are the numbers actually seen. Martin (1939) was unable to produce evidence of reduction division in *Chaetangium saccatum*, but suggests it occurs immediately after fertilization.

In contrast to *Scinaia furcellata*, species of the genera *Galaxaura* and *Actinotrichia* have three somatic phases. Kjellmann (1900) divided the genus into sections, based on certain structural differences of the cortex, but later Howe (1916, 1918) pointed out that all the forms of certain groups are tetrasporic and those of the others sexual. Moreover, a tetrasporic form of one group has its counterpart in, and often grew with, a form of the other group. After an examination of specimens of *Galaxaura obtusata* from Bermuda, Florida and the West Indies, Howe (1916) came to the conclusion that they were but the tetrasporophyte and gametophyte of the same species, showing the unusual feature of structural dimorphism. He later (1918) produced evidence of a similar nature for species belonging to other sections of the genus, and Børgesen (1927) also confirmed Howe's conclusions after examining specimens of *G. flagelliformis* from the Canary Islands. In contrast it should be noted that Harvey-Gibson & Knight (1913) report the occurrence of tetrahedrally divided tetrasporangia and cystocarps on the same specimen of *G. adriatica*, and Kjellmann (1900) found cruciately divided tetrasporangia and female sex organs on the same individual of *G. Giesingiana*. Cytological and experimental work would greatly add to the value of these observations.

Nothing is known of the nuclear history of species of the Neccariaceae, and although we have some cytological information about the neighbouring family, the Bonnemaisoniaceae, it is very inadequate. Svedelius (1933) while investigating the anatomical developments of *Asparagopsis armata* and *Bonnemaisonia asparagoides*, has also recorded some cytological findings. These, he postulates, show that reduction division takes place immediately after fertilization, in the hypogynous cell in the case of *Asparagopsis armata* (Fig. 1*d*) and in the carpogonium of *Bonnemaisonia asparagoides*. In both species four nuclei† result from the division of the nucleus of the carpogonium, two of these contributing to the development of the gonimoblast in the case of *Asparagopsis armata* and one only in *Bonnemaisonia asparagoides*. The evidence (as provided by the figures) in favour of the first division of the nucleus of the mature carpogonium being a reduction division is not very convincing, and there is little evidence that the chromosomes are paired, especially in the case of *B. asparagoides*. In these and other figures, chromatin threads are shown connecting the chromosomes. Kylin (1916*b*), in his investigation of *B. asparagoides*, could not be sure

* The division of the carposporangium of *Batrachospermum Breutelii* to form gemmae (Skuja, 1933) appears to be more in the nature of premature germination of the whole sporangium, not the spore, possibly influenced by external conditions.

† The presence of a tetrad of nuclei in the carpogonium cannot be taken as corroborative evidence of a reduction division as it is to be found in cases where there is no reduction division at that stage, e.g. *Spermothamnion Turneri* and *S. Snyderae*.

that reduction division takes place at the beginning of the cystocarp formation. He found twenty 'bodies' present in the prophase of the first nuclear division in the carpogonium—that is, the same number as in the somatic divisions—but could not ascertain whether they were single or double. Thus there appears to be little doubt that the same number of chromosomes is present in the nucleus of the carpogonium prior to any fertilization and in the cells of the gonimoblast, but whether there is a fusion of male and female nuclei and a subsequent reduction division seems uncertain. The whole question is further complicated by Feldmann & Mazoyer's (1937) and J. & G. Feldmann's (1939, 1939*a*) assertions, apparently well-founded, that the carpospores of *Asparagopsis armata* develop into *Falkenbergia rufolanosa* and those of *Bonnemaisonia asparagoides* into *Hymenoclonium serpens*. Both *Falkenbergia rufolanosa* and *Hymeonclonium serpens* bear tetrasporangia. Cytological information regarding these tetrasporic phases is lacking, but is highly essential for a correct understanding of the somatic and nuclear phases. J. & G. Feldmann (1939*a*) find that the carpospores of another species of *Bonnemaisonia*, *B. clavata*, germinate to give a plant very like *Hymenoclonium serpens*. Such marked dimorphism between a gametophyte and tetrasporophyte is new in our knowledge of this group.

The life history of another species of *Asparagopsis*, *A. hamifera*, is also far from clear. Records suggest that in recent years it has spread considerably, and Chemin (1929*a*), who considers Japan its centre of distribution, finds that both female and male plants are known in that locality, but in Europe and at Woods Hole, U.S.A., female plants bearing pseudocystocarps alone have been found. In this case it is possible that only one phase is present in localities away from Japan, the plant having become entirely dependent on vegetative means of reproduction. The possibility should be not overlooked that one phase of the life history of *A. armata* also is omitted in some localities.

It is obvious from this survey that there is considerable variety regarding the sequence of somatic phases in the Nemalionales, sometimes within the limits of the one family or even genus. Further work will show how much variety, if any, the nuclear phases show.

2. Order II. Gelidiales

This order contains but one family, the Gelidiaceae, containing, according to Smith (1938), about six genera. The life history of the species included shows three somatic phases, but the cystocarp develops directly from the carpogonium. The cytology of this family has been completely neglected.

3. Order III. Cryptonemiales

This large order, comprising eleven families (Kylin, 1937), in which are included about 650 species with very varying characteristics, has received almost no attention from cytologists. The feature common to all families of this order is the development of auxiliary cells on a special filament of the gametophyte. The only account dealing with cytology is that of Yamanouchi (1921) for *Corallina officinalis* var. *mediterranea*, showing that the sexual plants are haploid, with 24 chromosomes, and the carposporophyte and tetrasporic plants diploid, with 48 chromosomes. Fusion of a spermatial nucleus with the nucleus of the carpogonium has been observed and reduction division takes place in the tetrasporangium. This alga therefore has three somatic and two nuclear phases. On the basis of morphological evidence it is permissible to suppose that many, if not the majority, of the algae belonging to this order have a similar life history, but as long ago as 1917 Rosenvinge wrote: 'There are a considerable number of Cryptonemiales which differ with regard to the course of development from the typical diplobiontic forms', and accurate information regarding them is still needed. It is profitable to call attention again here to some of the facts to which he was referring. The occasional appearance of both sexual organs (or cystocarps) and tetrasporangia on the same plants has been recorded for the following species: *Dudresnaya coccinea* (Dumontiaceae), *Polyides rotundus* (Rhizophyllidaceae), *Cruoria pellita* and *Petrocelis Hennedyi* (Cruoriaceae), *Lithothamnion Lenormandii*, *L. flavescens* and *L. Strömfeltii*, and *Melobesia farinosa* (Corallinaceae).

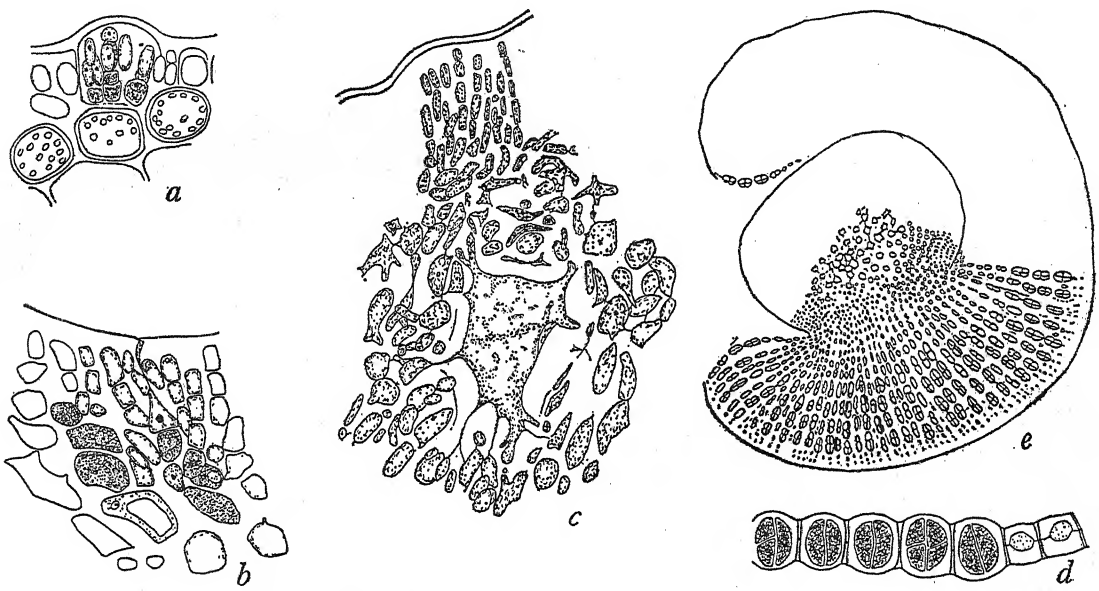
While tetrasporangia are unknown for *Gloiosiphonia capillaris* (Gloiosiphoniaceae) in France and Denmark they have been reported from Norway and Sweden, and although tetrasporic plants are known for *Polyides rotundus* (Rhizophyllidaceae) they are not nearly so frequent as the sexual plants. Sexual plants only are known for *Acrosymphyton purpurifera*. In contrast, tetrasporic plants only are known for the following: *Hildenbrantia* spp., *Cruoriopsis gracilis*, *Rhododermis Georgii* (Squamariaceae), and species of *Lithothamnion* (Corallinaceae). *Rhododermis elegans* was thought at one time to have no sexual plants, but later some male specimens were found in north-east Greenland. In *Petrocelis Hennedyi* Rosenvinge also suggests that the gonimoblast develops parthenogenetically. Detailed cytological investigations with very extensive field observations throughout the growth season are very necessary in the case of such algae.

4. Order IV. Gigartinales

This order is characterized by the fact that the gonimoblast develops not from the carpogonium direct but from an auxiliary cell which is a vegetative cell of the gametophyte, the details concerning it varying from family to family. Of the nineteen families (Kylin, 1937) in this large order comprising about 500 species, none has received any

but the most scanty cytological investigation. Apart from the desirability of accounts of the nuclear history in some of the common forms, e.g. *Chondrus crispus*, where there are the usual three somatic phases, cytological information is much needed in the case of other species showing unusual features. Comparatively recently, observations have been made showing that some species, belonging to the Phyllophoraceae and Nemastomaceae, have two somatic phases only, the gametophyte and a tetrasporic nemathecium (Fig. 2*e*), originating from the auxiliary cell. These nemathecia had previously been considered epiphytic parasites. Phillips (1925) gives a general account of the facts known at that time for the genera *Phyllophora*, *Gymnogongrus* and *Ahnfeldtia*, and although some progress has been made since then much remains in doubt. Phillips points out that where no 'parasites' are known the plants have both cystocarps and tetrasporangia. 'Parasites' have tetrasporangia only, and the 'host' has either cystocarps only or no reproductive organs at all. Since Phillips wrote it has been ascertained that the tetraspore-bearing structure previously known as *Actinococcus subcutaneus* develops from the auxiliary cell of *Phyllophora Brodiaei* (Fig. 2*c*), but the events immediately preceding this appear to be in need of further clarification. For example, the carpogonia (Fig. 2*b*) often look degenerate and may be without a trichogyne (Rosenvinge, 1929), a high percentage aborting (Kylin, 1930). The carpogonium unites with the auxiliary cell (Kylin, 1930) and the tetraspores from the resulting nemathecium germinate into plants, considered to be young *P. Brodiaei* (Rosenvinge, 1929). The cytological evidence produced by Claussen (1929), on the basis of which he declared that there is a reduction division in the tetrasporangium, cannot by any means be taken as conclusive. Further research of both a cytological and morphological nature to give essential information about the life history of this species is obviously needed. It should be noted that another species of the genus, *P. membranifolia*, has three somatic phases, the tetrasporangia being produced in nemathecia.

This dissimilarity of life history within the limits of a genus is found also in *Gymnogongrus*. Phillips (1925) and Doubt (1935) point out that the thirty or more species in the genus can be divided into those without sexual reproductive organs but with the so-called 'parasite', and those with male and female organs and no 'parasite'. A species in one group may have its counterpart in the other group, so alike as to be almost indistinguishable but with a different geographical distribution, e.g. *G. linearis* and *G. platyphyllus* form one pair (Doubt, 1935) and *G. pusillus* and *G. Griffithsiae* another (Feldmann & Mazoyer, 1938). The first species of these pairs have normal internal cystocarps, the others external tetrasporic nemathecia. Doubt found both male and female plants of *G. linearis* and evidence of fertilization followed by the production of carpospores, but the same chromosome number, namely, 6, occurs in



Text-fig. 2. *Phyllophora Brodiaei* (Rosenvinge). *a*, young spermatangial crypt (×625); *b*, two procarps, one without trichogyne (×390); *c*, auxiliary cell and nemathecium initials (×390); *d*, tetrasporangial filament (×625); *e*, *Gymnogongrus Griffithsiae* (from Kylin), mature nemathecium.

both the vegetative cells and the germinating carpospores. Female plants only of *G. platyphyllus* were found, the carpogonia appearing normal. The so-called nemathecial 'parasite', *Actinococcus chiton*, develops from the auxiliary cell and produces tetrasporangia. No evidence of fertilization has been found and the chromosome number, 8, is the same in the vegetative cells, the carpogonial branch and the germinating tetraspores. Diakinesis has not been seen in the tetrasporangium. Gregory (1934) found essentially the same series of events to be true for *G. Griffithsiae*. *G. norvegicus* is unlike the other species of the genus in that plants with normal cystocarps as well as plants bearing tetrasporic nemathecia (previously known as *Actinococcus peltaeformis*) occur. Chemin (1927, 1929) found the germination of the tetraspores and carpospores to be strikingly similar, and the young thalli formed from the tetraspores have the structure of *Gymnogongrus*.

More recently, Smith & Hollenberg (1943) have stated that they consider the life history of *Schizymenia epiphytica* (Nemastomaceae) to be of the same type as that of *Phyllophora Brodiaea*. The life history of *Ahnfeldtia plicata*, still incompletely known, appears to be of a 'derived' type. Male plants have been described (Gregory, 1934), but there is no indication of any procarpic structure. Prior to the formation of the nemathecia the cortex hypertrophies and then cells with deeply staining contents and pointed apices—but with function unknown—make their appearance amongst the cells of the cortex, originating from the medulla. Subsequently 'generative' cells (Rosenvinge, 1931) are to be seen at the surface of the cortex and from these develop the monosporangia-bearing nemathecia. On germination of the monospores Rosenvinge (1931) and Chemin (1930) both obtained disks which they consider to be young *Ahnfeldtia plicata*. Rosenvinge (1931) states that there is no sign of reduction division in the monosporangia and that indications of four chromosomes in the sporangium and vegetative cells need confirmation.

As in the preceding order so in this one there are reports of the occurrence of tetrasporangia and sexual organs (or cystocarps) on the same specimens, e.g. *Platoma Bairdii* (Nemastomaceae), *Agardhiella tenera*, *Solieria chordales* (Solieriaceae), *Catenella Opuntia* (Rhabdoniaceae), *Hypnea Valentiae* (Hypnaceae), *Gracilaria confervoides* (Gracilariaceae), and *Phyllophora membranifolia* (Phyllophoraceae). Tetrasporangia are unknown for the genus *Turnerella* (Solieriaceae) (Taylor, 1937) as well as for *Gigartina mamillosa* (Gigartinaceae). However, Newton (1931) figures tetrasporangia for the latter species. The case of *Halarachnion ligulatum* (Furcellariaceae) is of special interest, for although tetraspores are unknown in nature, Dammann (1930–2) germinated the carpospores, and after five months, when only $\frac{1}{2}$ mm. high, the germlings produced and matured tetrasporangia.

Platoma Bairdii is the best known example of parthenogenesis in the Florideae. Only female sexual organs are known but the cystocarps develop normally, the nuclear condition being unknown. It is possible that *Furcellaria fastigiata* (Furcellariaceae) is parthenogenetic occasionally, if not always.

5. Order V. Rhodymeniales

There are two families in this order, the Rhodymeniaceae and the Champiaceae containing about 130 species (Smith, 1938). There is no cytological evidence, but the morphological evidence supports the view that the majority of the Rhodymeniaceae are of the *Polysiphonia* type. *Rhodymenia palmata* (one of the commonest red algae) and *Halosaccion ramentaceum** are noteworthy since only male and tetrasporic plants are known. Westbrook (1928) has examined the various stages of nuclear division in the tetrasporangium of the former species and has seen a condition which appears to be synizesis, but diakinesis has not been observed. She inclines to the view that the nuclear divisions in the tetrasporangium are meiotic however, and that female plants will be found.

Similarly in the Champiaceae we find only slight pieces of cytological evidence available. No doubt many life histories are of the *Polysiphonia* type. This appears to be the case in *Lomentaria clavellosa* (Svedelius, 1937), but it seems possible that there is no reduction division in the tetrasporangium of the nearly related *L. rosea*, a species for which sexual organs are unknown (unless a Japanese record (Segawa, 1936) is accepted). Svedelius (1937) describes and figures abnormal cytological features, not of a kind to be associated necessarily with the absence of a meiotic division, in the tetrasporangium of this species. For example, 'bodies' inside the nucleolus are said to be chromosomes and they are figured dividing in the nucleolus also. As some of the figures show considerable contraction of the cytoplasm I question whether the appearances figured are normal or whether the investigator saw the later stages of nuclear division. A normal early prophase for the tetrasporangium mother cell and a normal late prophase in mature tetraspores are figured, but the intermediate stages need confirmation.

Polysporangia occasionally replace tetrasporangia in *Gastroclonium ovale* and *G. Goulteri*, and this type of sporangium is the only one known in two species of *Coeloseira* on the Pacific Coast of North America. In the latter genus Hollenberg (1940) describes the initials as being multinucleate. One of the nuclei enlarges and divides equationally to give four, all of which the investigator thinks undergo reduction division (contrast with *Spermothamnion Snyderae*). It should be noted that the final number of spores

* Taylor (1937) reports tetrasporangia only for the north-east coast of North America.

varies from 12 to 16, whereas 16 is the expected number should the divisions take place as stated. Spermatangia are unknown in the genus and cystocarps are known for only one of the two species. In *Bindera saccata* and *Chylocladia kalifomis* tetrasporangia and sexual organs occur on the same plant.

6. Order VI. Ceramiales

This order contains the greatest number of species (about 900; Smith, 1938) but is divided into only four families, the Ceramiaceae, Delesseriaceae, Dasyaceae and Rhodomelaceae, all characterized by the fact that the auxiliary cell, from which the gonimoblast develops, is not formed until after fertilization. Some of the most thorough and convincing cytological investigations concern species of this order; the first of these was Yamanouchi's researches (1906) on *Polysiphonia violacea*.

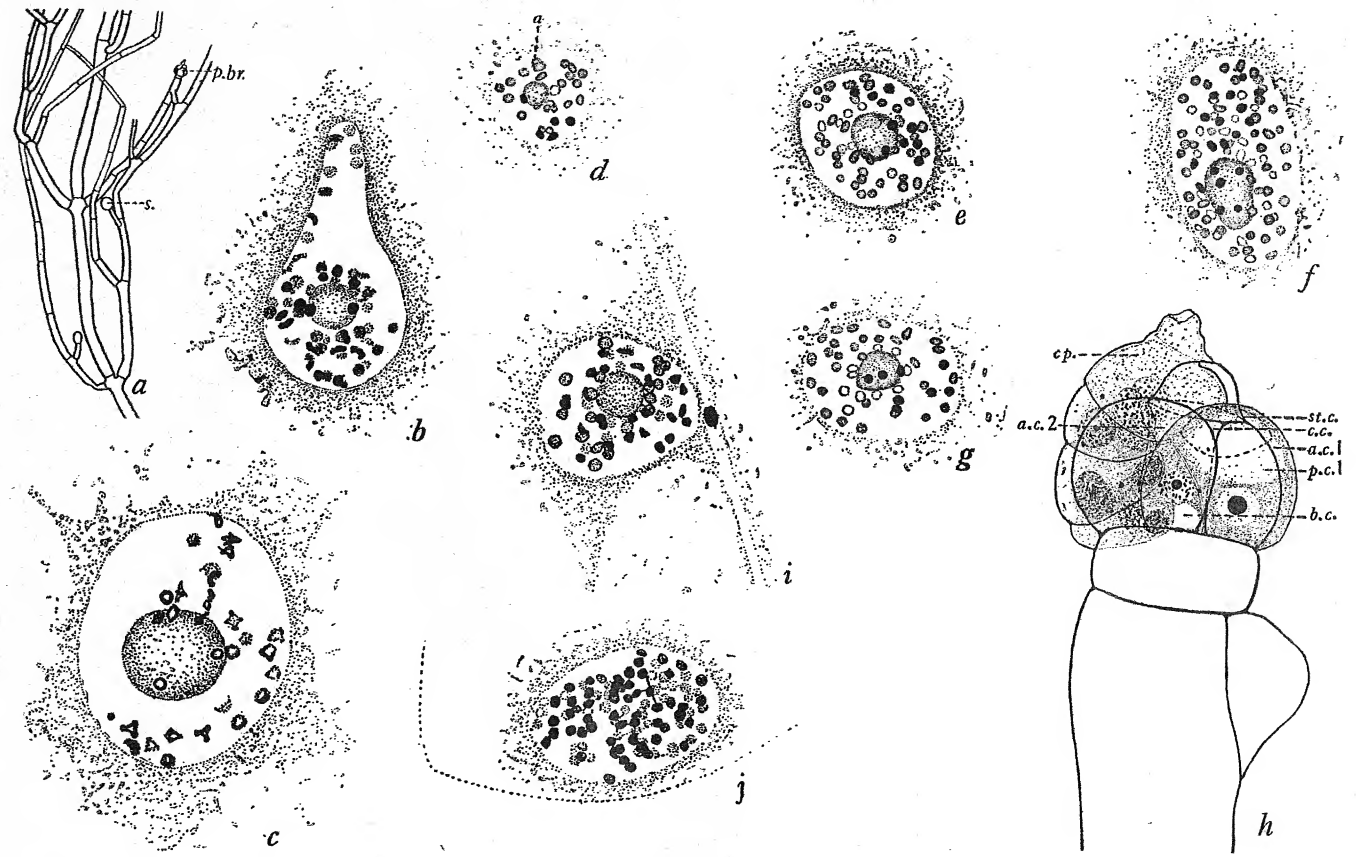
The majority of the Ceramiaceae are undoubtedly cytologically diphasic with three somatic phases, as was found to be the case in *Griffithsia Bornetiana* (Lewis, 1909). On account of considerable peculiarities in the method of nuclear division the storing of chromosomes in the nucleolus and the absence of a spireme stage, Kylin (1916*a*) thought a repetition of Lewis's investigations was desirable. He used the nearly related *Griffithsia corallina* and found that the method of nuclear division is similar to that usually described for the Ceramiales and unlike that described by Lewis. Fertilization takes place and there is a typical meiosis in the tetrasporangium. The haploid chromosome number is 20, that given by Lewis being 7. Both Lewis and Kylin, it is worth noting, found that the tetrasporic outnumber the sexual plants, and Rosenvinge (1923) has reported the same in Denmark for *Ceramium* spp. and *Antithamnion plumula*. More recently Mathais (1928) on investigating *Callithamnion brachiatum* has found 10 chromosomes in the nuclei of the sexual plants and 20 in the nuclei of the tetrasporic plants, reduction division accompanying the formation of the tetraspores. There has been some criticism (Westbrook, 1930, 1935), probably justifiably, of Mathais's interpretation of the appearance of the nucleus. It would seem that some of his other statements are likewise open to question. For example, cystocarps were found on a plant bearing tetrasporangia, and the 'supporting filament' is stated to be diploid, the carpospores haploid and the cytology of the tetrasporangium normal. Of these cells only the haploid carpospores are figured, one of the illustrations being a photomicrograph in which it is impossible to distinguish any nuclear details. To explain the haploid state of the carpospores Mathais postulates either a double reduction division or apogamy with a single reduction division. As both of these explanations are so anomalous they cannot be accepted without further evidence.

In *Spermothamnion Turneri* tetrasporangia, sexual organs (particularly the female) and cystocarps commonly occur on the same plant and often in close proximity* (Fig. 3*a*). This species has been thoroughly examined (Drew, 1934 and 1943) on the British coast and at Woods Holé on the east coast of North America. Along the British coast haploid (30), diploid and triploid plants occur as well as individuals which have approximately half the triploid number of chromosomes (Fig. 3*d–g*). The same is true for the American coast in the neighbourhood of Woods Hole, with the exception that no haploid plants have been found there in spite of indirect evidence of their existence. British haploid plants bear sexual organs and cystocarps and occasionally undivided sporangium mother cells, which may be either in a healthy condition or degenerated. Diploid plants, in particular those in American waters, are characterized by the occurrence of both functional sexual organs and normal tetrasporangia (Fig. 3*a*). A normal reduction division precedes the formation of the tetraspores (Fig. 3*c*). The procarps develop normally and no reduction division precedes the formation of the carpogonium, which is diploid (Fig. 3*b*). Fertilization has not been observed but diploid American plants are known to produce cystocarps, the divided carpogonium nucleus and nuclei of the gonimoblasts being tetraploid (Fig. 3 *h, j*). Two triploid gonimoblasts, one British, one American, have also been seen, presumably the result of union of haploid and diploid sexual cells. All triploid plants examined have been sterile (with the exception of one bearing a degenerated procarpic branch), but the existence of plants with half the triploid number of chromosomes suggests the formation of tetrasporangia and a reduction division.† These results disagree with those given by Schussnig & Odle (1927) for *S. roseolum* (probably not specifically distinct from *S. Turneri*), but as no vegetative nuclei were seen in division by these investigators they had no means of deciding which plants were haploid and which diploid. Without this their conclusions that reduction division takes place in the tetrasporangium on tetrasporic plants only, and not in the tetrasporangium of sexual plants, carry no weight. Other cytological evidence is lacking. The occurrence of sexual organs and tetrasporangia on the same plant is a common event in this family. Kniep (1928) lists twenty-eight species in which it is known to occur, many of them belonging to the genus *Callithamnion*.

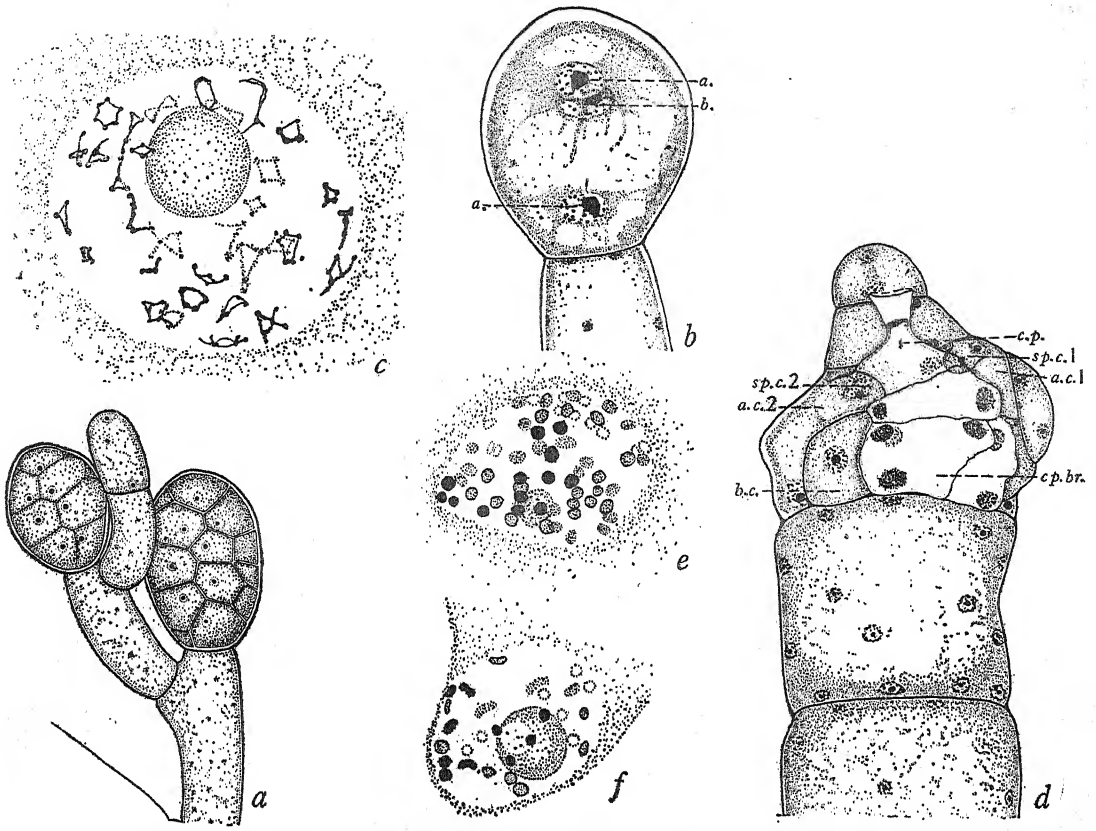
Some members of the Ceramiaceae have polysporangia, containing several spores; these develop as well as, or in place of, tetrasporangia, *Spermo-*

* In *Callithamnion tetricum* (Westbrook, 1930*a*) the supporting cell may give rise to both a tetrasporangium and a carpogonial branch.

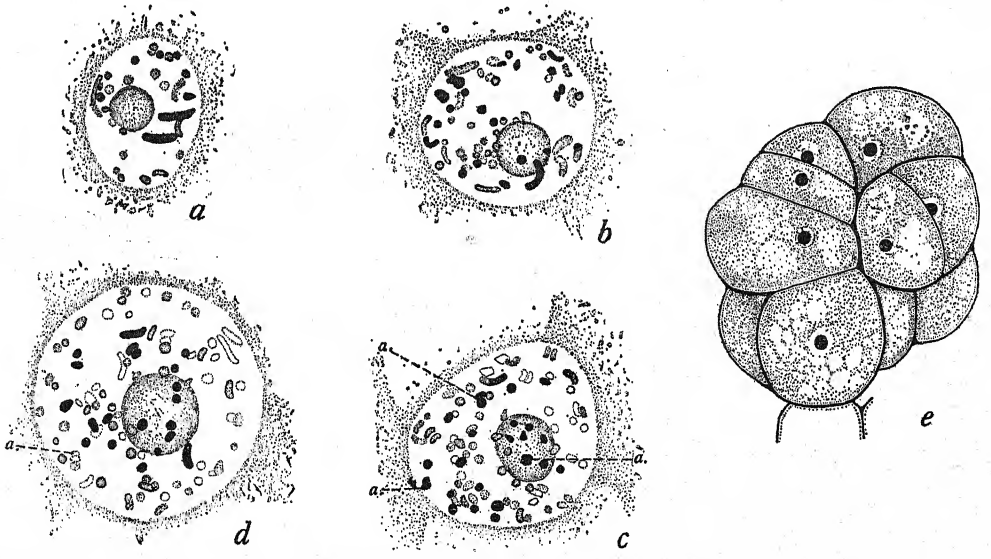
† For a diagrammatic summary of the information regarding *S. Turneri*, the reader is referred to *Ann. Bot., Lond.*, 7, 1943.



Text-fig. 3. *Spermothamnion Turneri* (Drew). *a*, procarpic branch and young sporangium on same filament (× 59); *b*, diploid carpogonium nucleus of same procarpic branch, and *c*, nucleus of sporangium with haploid number of bivalent chromosomes (both × 2350); *d*, haploid; *e*, diploid; and *f*, triploid somatic nuclei (× 2150); *g*, nucleus with 48 chromosomes (× 2150); *h*, procarp showing post-fertilization changes (× 600); *i*, diploid nucleus of basal cell of this procarp; and *j*, tetraploid nucleus of its carpogonium (both × 2350). Chromosomes behind nucleolus or other chromosomes not figured.



Text-fig. 4. *Spermothamnion Snyderae* (Drew). *a*, polysporangia (× 188); *b*, three nuclei in young polysporangium in late diakinesis (×250); *c*, early diakinesis (×2120); *d*, post-fertilization changes in procarp (×485); *e*, diploid nucleus of sporogenous cell (*sp.c.*2); and *f*, haploid nucleus of auxiliary cell (*a.c.*2) of same procarp (both ×2120).



Text-fig. 5. *Plumaria elegans* (Drew). Nucleus of apical cell of *a*, male plant (haploid); *b*, tetrasporic plant (diploid); and *c*, plant with paraspores (triploid); *d*, triploid nucleus of parasporangium; *e*, mature parasporangium (*a–d* × 2150, *e* ×625). Chromosomes behind nucleolus or other chromosomes not figured.

thamnion Turneri being an example of the former and *S. Snyderae*, *Pleonosporium Vancouverianum* and *P. squarrosum* of the latter. A detailed examination of *Spermothamnion Snyderae* (Drew, 1937) has shown that the sporangium initial may contain from 2 to 9 nuclei (Fig. 4*b*), and they all undergo a reduction division (Fig. 4*c*) into four and the spores are uninucleate (Fig. 4*a*). Polyspores, in this case, are therefore homologous with tetraspores, the polysporangium representing a compound tetrasporangium. The sexual plants and reproductive cells have nuclei with 32 chromosomes and the sporangium-bearing plants 64. Fertilization has not been observed, but in one specimen (Fig. 4*d*) two nuclei in the carpogonium are to be seen in the process of degenerating, while each of the two uninucleate sporogenous cells is in open connexion with its auxiliary cells. One of these sporogenous cell nuclei is in division and has the diploid number of chromosomes (Fig. 4*e*), whereas the nucleus of its auxiliary cell, in the same condition, has the haploid number (Fig. 4*f*). Drew (1939) has shown that the parasporangium of *Plumaria elegans* (Fig. 5*e*) is very different both morphologically and cytologically from the polysporangium just described. Normal haploid (31) gametophytes (Fig. 5*a*), diploid carposporophytes and diploid tetrasporophytes (Fig. 5*b*) occur, but the triploid (Fig. 5*c*) is by far the most common phase as well as the most widely distributed, being able to exist in colder waters. Parasporangia develop on triploid plants only, and the paraspores are formed without any reduction in the chromosome number (Fig. 5*d*). They reproduce the triploid, and the fact that they are formed in great abundance accounts in part no doubt for the preponderance of the triploid. Occasional tetrasporangia develop on the triploid, but their chromosome complement as well as their fate is unknown*. The cytology of the haploid and diploid plants as well as the reproductive organs is normal. Other species for which parasporangia, so-called, are known include *Callithamnion Hookeri*, *Ceramium diaphanum*, *C. strictum*, *C. Deslongchampsii*, *C. vertebrale* and *Antithamnion plumula*. Cytological investigations of these forms are awaited. The monospores of *Monospora pedicellata* and the seirospores of *Seirospora Griffithsiana* are well known, but there is no information concerning the nuclei of these spores. As in other families, species occur for which tetrasporic plants only are known, namely, *Callithamnion tripinnatum*, *Antithamnionella sarniensis*, *Ceramium callipterum* and *Callithamnionella tingitana*.

Coming to the Delesseriaceae, *Delesseria sanguinea* received very full attention from Svedelius (1911, 1912 and 1914) who showed that the life history has three somatic phases. As a result of fertilization of the carpogonium, which was observed, the carposporophyte is diploid. Reduction division takes place in the tetrasporangium. The haploid chromosomes number is said to be 20, based on counts of the prophase when fine threads still connect the chromosomes. From the information available it seems probable that the other algae of this family have a similar life history. Svedelius's (1914 *a*, *b*) accounts of the nuclear history in the tetrasporangium of *Nitophyllum punctatum*, as well as in the sporangia formed on sexual plants of the same species, are of particular interest. The tetrasporangium initials on the tetrasporophyte are multinucleate, and after a period during which the nuclei divide equationally to give a maximum total of 12, one nucleus takes up a central position and undergoes a reduction division (Fig. 6*a*). The remainder degenerate before the meiotic division is complete. The sporangium initials formed on a sexual plant are likewise multinucleate and equational divisions take place. Then, however, one nucleus enlarges (Fig. 6*b*) and the rest degenerate, but the surviving one does not divide and the liberated spore is uninucleate. Mention should be made of the polysporangia of *Gonimophyllum Skottsbergii* in which Kylin (1924) has shown that the initial is uninucleate but successive divisions result in 30–50 nuclei. No details regarding the nuclear divisions are available.

In the very small family, the Dasyaceae, Westbrook (1935) has investigated the early stages in the division of the nucleus of the tetrasporangium of both *Dasya arbuscula* and *D. ocellata*. The former shows the typical appearances of a reduction division but not the latter. This is of particular interest, for whereas sexual as well as tetrasporic plants of the former occur only tetrasporic plants of *D. ocellata* are known. Further more detailed investigations should be of significance.

To the Rhodomelaceae belongs *Polysiphonia violacea*,† the first red alga to be the subject of a thorough cytological investigation. Yamanouchi's account (1906) will remain a classic because of its completeness and the fact that the significance of

* Estimating the number of chromosomes in the nuclei of these algae is not easy but a reasonable degree of accuracy can be attained. I have used whole mounts and counted chromosomes in the late prophase, when scattered throughout the nuclear area. The chromosomes of each successive focal plane have been mapped, and even chromosomes behind the nucleolus, unless very near it, have been clearly visible. In *S. Turneri* 125 nuclei were thus mapped in detail, 64 in *S. Snyderae* and 78 in *P. elegans*. Others were estimated more roughly, but only clear examples were used. Of the haploid counts made in detail in *P. elegans* 65 % gave 31 and 98 % numbers between 30 and 32; of the diploid counts 42 % gave 62 and 92 % numbers between 60 and 64, and of the triploid counts 20 % gave 93 and 85 % numbers between 87 and 99. It is not possible to attain the same degree of accuracy with the higher numbers, but there has been no possibility of confusing the haploids, diploids and triploids.

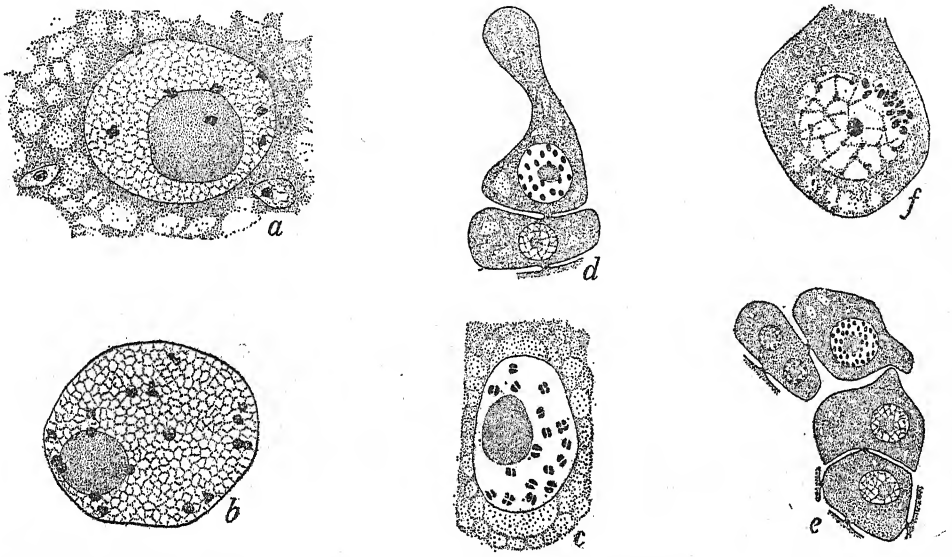
† Since reported by Taylor (1937) to be *P. flexicaulis* (Harvey) Collins. The genuine *P. violacea* is European.

the cystocarp and the tetraspore-bearing plants was revealed for the first time. It has become the 'type' of one variety of life history in this group of plants. The sexual plants were found to have 20 chromosomes, as well as the sexual cells themselves. The fusion of the haploid male and female nuclei was observed (Fig. 6*f*). The nucleus divides once in the carpogonium before entering the auxiliary cell, and in this division 40 chromosomes are to be seen as well as in the young carposporangia. There are 40 chromosomes also in the nuclei of the tetrasporic plants, and reduction division in the tetrasporangium mother cell with the subsequent production of haploid tetraspores was followed in great detail. Sporangia occasionally develop on sexual plants, but

Laurencia pinnatifida and *L. hybrida*, *Rhodomela subfusca*, *Brogniartella byssoides*, *Chondria tenuissima* and *C. dasyphylla*. *Trailliella intricata*, sometimes included in this family, presents a problem awaiting investigation, as only tetrasporic plants are known. Kniep (1928) lists ten members of this family as plants where sexual organs and tetrasporangia occur on the same specimens. These include species of *Polysiphonia*, *Pterosiphonia bipinnata*, *Laurencia hybrida* and *Brogniartella byssoides*.

III. DISCUSSION

As might perhaps have been expected, more is known about the somatic than the nuclear phases in



Text-fig. 6. *Nitophyllum punctatum* (× 1500, Svedelius). *a*, diakinesis of tetrasporangium mother nucleus; *b*, nucleus of sporangium on sexual plant, *Rhodomela virgata* (Kylin); *c*, diakinesis of tetrasporangium mother nucleus; prophase in carpogonium *d*, during last division prior to; and *e*, first division after fertilization; *f*, *Polysiphonia violacea* (Yamanouchi). Fertilization in carpogonium.

in almost all the cases investigated the nucleus had not completed the first divisions although invagination had started. The cleavage furrows were never seen to reach the centre of the cell. Kylin's work on *Rhodomela virgata* (1914) was a very thorough confirmation of the broad outline, and many of the details of the work on *Polysiphonia violacea*, and it provides one of the best accounts of such a type of life history. The haploid chromosome number is 20, and reduction division in the tetrasporangium, including diakinesis (Fig. 6*c*), is figured very clearly in all stages. Fertilization was not seen, but 20 chromosomes were counted in the prophase stage of the carpogonium nucleus (Fig. 6*d*) and 40 during the first division of the zygote nucleus (Fig. 6*e*).

Westbrook (1935) has seen diakinesis in the sporangia of some other members of the family:

the Florideae, almost all of the cytological work having centred on species belonging to the Nemalionales and Ceramiales. As these two orders occupy extreme systematic positions, being at either end of the generally accepted classification, an almost untouched field awaits exploration in the Gelidiales, Cryptonemiales, Gigartinales and Rhodymeniales. This being the situation it is clear that caution should be exercised in making generalizations regarding the Florideae as a group, and on occasions judgement should be suspended until further data are available. While our information is still too limited to segregate the investigated species into 'typical' and 'atypical', the morphological evidence suggests that the *Polysiphonia* type of life history, with three somatic and two nuclear phases, is the most general throughout the Florideae; among

species without tetraspores the *Nemalion* type with two morphologically dissimilar haploid phases, separated by a very short nuclear diplophase, may be regarded as probably representative. This was the justification for Svedelius's subdivision of the Florideae into diplo- and haplobiontic. Nevertheless, the desire to achieve a generalization must not be allowed to mask the important fact that there exist many species which do not conform either morphologically or cytologically to these types. Such species can be classified into the following groups, but as the information available is scanty the arrangement of the ten groups is somewhat arbitrary. Groups 1–4 show possible affinity with the *Polysiphonia* type, as does 5 with the *Nemalion* type; 9*b* and 10 are probably 'abnormals' rather than representatives of new types.

(1) Species with sexual organs and tetrasporangia on the same individuals. This occurrence seems to be general for some species and exceptional for others. The majority of the recorded cases (about 70) concern species of the Ceramiales, but this phenomenon is known throughout the Florideae.

(2) Species with polysporangia and no tetrasporangia.

(3) Species with parasporangia in addition to tetrasporangia.

(4) Species with a special type of spore, such as the seirospores of *Seirospora Griffithsiana* and the monospores of *Monospora pedicellata*.

(5) Species in which tetrasporangia replace carposporangia.

(6) Species for which morphologically dissimilar but cytologically similar gametophytes and tetrasporophytes are said to exist.

(7) Species in which a tetrasporic or monosporic nemathecium results from what appears to be an actual or substitute fertilization.

(8) Species for which tetrasporic specimens only are known.

(9) (*a*) Species such as *Rhodochorton violaceum* with three somatic phases, one producing tetrasporangia, in a genus of otherwise very simple forms, the majority reproducing by means of monosporangia only. (*b*) Species without tetrasporangia belonging to genera the other species of which normally produce tetrasporangia.

(10) Species for which either no male or female plants are known.

In groups (1), (2) and (3), one example of each has been investigated cytologically by me. These researches produced new and unexpected facts including the existence of polyploidy, the presence of fertile and functional sexual organs on diploid plants, the existence of a compound tetrasporangium, the polysporangium, and the association of a special type of sporangium with triploid plants. The development of functional sexual organs on diploid plants provides one explanation of the occurrence of sexual organs, cystocarps and tetrasporangia on the same specimens. This is so common that not many years after Yamanouchi's classical account of the life history of *Polysiphonia violacea* (1906), Harvey-Gibson & Knight (1913) expressed the opinion that they could not all be abnormalities and later writers have supported this view. The development of the sexual organs on diploid as well as on haploid plants provides automatically the possibility of polyploidy, a state which may be more common in the Florideae than the records suggest.

Questions of some importance are raised by those species which do not conform to the *Nemalion* and *Polysiphonia* types. Thus *Phyllophora Brodiaei* (7) and similar forms raise the question of the general homology between the product of fertilization (either actual or substitute) in these species, and the cystocarp of *P. membranifolia* on the one hand and forms such as *Liagora tetrasporifera* on the other. Another question of homology is raised by the tetrasporophytes of *Asparagopsis armata* and *Bonnemaisonia asparagoides* (6), whether they are shown to be haploid or diploid. The discovery of the place of reduction division in *Liagora tetrasporifera* (5) will show either that the diploid nuclear phase may be of varying length, even in one genus, or that reduction division is not always associated with the formation of the tetraspores. This latter possibility has been envisaged by various workers, but there is still no absolutely irrefutable cytological evidence that this is so.

The preceding summary shows how incomplete and even scanty is our knowledge of this group. Nevertheless, in the floridean literature, along with the facts, there is a wealth of speculation regarding the origin and development of the *Nemalion* and *Polysiphonia* types of life histories, with attempts to 'relate' the more recently discovered 'abnormalities' to the same general scheme. A composite summary of the most generally accepted of these suppositions would give a picture of the primitive red alga as one where reduction division takes place immediately after fertilization, the fertilized egg cell dividing up within its own wall. Later the cystocarp was 'intercalated' to give the *Nemalion* type as we know it to-day and then by a 'postponement of reduction division', either sudden or gradual, the *Polysiphonia* type appeared. With by no means all the necessary information forms such as *Liagora tetrasporifera* and *Helminthocladia Hudsoni* are looked on as intermediate types. Then the life history having become more and more elaborate the process is held to have been reversed and 'reduced' types of life history appear. Among these may be quoted such species as *Lomentaria rosea*, where the sexual phase has been 'dropped out' and only tetrasporic plants are left, and *Phyllophora Brodaiei* where the product of fertilization, either actual or substitute, is supposed to represent the tetrasporic phase and not the cystocarp. It may be that further research will show that in some instances the sequence of events indicated above is substantially correct, but meanwhile it is clearly a case for a suspension of judgement with the rider that such speculations may even impede progress in that they give the work so far accomplished a spurious air of completeness. It is

not too much to say that in several of these instances which have been most productive of speculation a few critical nuclear observations would serve to clarify the position and afford the required refutation or substantiation. Even if the assumption is made that all the published floridean cytology is up to the best contemporary standards*—which seems unlikely—the total body of such work is still insufficient for inductive treatment. Indeed, no real progress will be made until cytological investigation together with experimental work on questions of nutrition and the effects of environment are undertaken on an adequate basis.

Broadly speaking three lines of cytological research are indicated: (1) comprehensive studies of members of the orders Gelidiales, Cryptonemiales and Gigartinales which have received almost no attention; (2) critical examination of species which do not conform to the *Nemalion* and *Polysiphonia* types and listed above; (3) a repetition and extension of the work already done on the so-called representative types. In any of the work suggested ecological information relevant to the whole range of distribution of a species is desirable, together with statistical data on the occurrence of abnormal individuals, e.g. plants bearing both sexual organs and tetrasporangia. Although as a group the Florideae is in many ways unique yet at present there is a striking absence of specific physiological concepts, while the relation of physiological factors to growth, form and reproductive processes is almost completely unexplored. A comparative study of the methods of nutrition of the varying types of cystocarp, to mention one example only, would obviously be of great value, but this is looking far ahead.

Halarachnion ligulatum may be cited in illustration of the need for pure culture investigations, for whereas tetrasporic plants are unknown in nature, Dammann (1930–2) was able to germinate the carpospores in Berlin, and after five months, when only ½ mm. high, the germlings produced and matured tetrasporangia. Possibly tetrasporophytes do exist in the sea, but if they do not their absence needs explanation.

In conclusion, the importance of critical technique in the preparation of material and the use of first-class lighting and optical equipment should be stressed. Without these, results cannot be relied upon and may only add to the confusion already existing. Above all, it is essential to maintain a much more critical attitude to these problems than has been obvious in the past and to withhold 'explanations' or suspend judgement until the necessary information has been obtained.

* The earlier view that the haploid chromosome number is uniformly 10 or 20 seems open to question.

IV. SUMMARY

(1) The need for a considerable extension of morphological and cytological work in the Florideae is pointed out. (2) Before proceeding to the survey current terminology is reviewed; some terms are discarded and definitions are given for those retained. The survey shows that our cytological knowledge is limited almost entirely to two of the six orders, the Nemalionales and Ceramiales. Members of the Nemalionales show great variety as regards the somatic phase sequence in the life history. Species investigated cytologically, to varying degrees of detail, are all of the *Nemalion* type, that is, two morphologically dissimilar haploid phases separated by a very short diplophase of the nucleus. The Gelidiales have been completely neglected, and apart from *Corallina officinalis* var. *mediterranea* so have the Cryptonemiales. In this latter order there are indications of 'atypical' phase sequence. This is also true of the Gigartinales which include in addition forms showing the *Phyllophora Brodiaei* type of life history, i.e. two morphologically dissimilar phases, one a gametophyte and the other a tetrasporic nemathecium developed from the auxiliary cell. The nuclear phase sequence is unknown. Information about the nuclear history in the Rhodymeniales is only sufficient to indicate problems in need of investigation. In the Ceramiales the *Polysiphonia* type of life history with three somatic phases (two morphologically similar) and two nuclear phases appears to be general, and detailed accounts of such exist for several species. Other types are known and representatives have been investigated: *Spermothamnion Turneri* has been shown to be cytologically polyphasic and to be characterized by the formation of functional sexual organs on the diploid; the polysporangium of *S. Snyderae* has been shown to be a compound tetrasporangium, and *Plumaria elegans* has been found to have three nuclear phases, the triploid bearing a special type of sporangium. In addition, some species show unusual somatic phases sequences, the nuclear phases being unknown. (3) Three lines of cytological research are envisaged associated with ecological and physiological investigations. (4) There is a striking absence of physiological concepts concerning the Florideae and hence the need for pure culture and general physiological investigations.

I wish to express my thanks to Prof. C. W. Wardlaw for his stimulating interest and help in the preparation of this article, and also to the Council of the University of Manchester for library facilities. Thanks are also due to Mr E. Ashby for his help in the production of the figures. Owing to war-time restrictions direct copying of previously published figures has not been possible. Instead they have been photographed and magnified considerably in order that the photographic image could be meticulously picked out in Indian ink. Subsequently the diagram was cleared by bleaching and reduced to its original size. I trust that faithful copies of the originals, as published, have thus been obtained.

V. REFERENCES

BAUCH, R. (1937): *Planta*, **26**, 365.
BØRGESEN, F. (1927): *K. danske Vidensk. Selsk.*, Biol. Medd., **6**, 3.
CHEMIN, E. (1927): *Bull. Soc. bot. Fr.* **74**, 912. — (1929): **76**, 305. — (1929*a*): *Rev. algol.* **4**, 29. — (1930): *Bull. Soc. bot. Fr.* **77**, 342.
CLAUSSEN, H. (1929): *Ber. dtsch. bot. Ges.* **47**, 544.
CLELAND, R. E. (1919): *Ann. Bot., Lond.*, **33**, 323.
DAMMANN, H. (1930–2): *Wiss. Meeresuntersuch.* **18**, 26.
DOUBT, D. G. (1935): *Amer. J. Bot.* **22**, 294.
DREW, K. M. (1928): *Univ. Calif. Publ. Bot.* **14**, 139. — (1934): *Ann. Bot., Lond.*, **48**, 549. — (1935): **49**, 439. — (1937): N.S. **1**, 463. — (1939): **3**, 347. — (1943): **7**, 23.
FELDMANN, J. (1939): *Bull. Soc. Hist. nat. Afr. N.* **30**, 87.
FELDMANN, J. & G. (1939): *C.R. Acad. Sci., Paris*, **208**, 1240. — (1939*a*): **208**, 1425.
FELDMANN, J. & HAMEL, G. (1934): *Rev. gén. Bot.* **46**, 528.
FELDMANN, J. & MAZOYER, G. (1937): *C.R. Acad. Sci., Paris*, **205**, 1084. — (1938): *Bull. Soc. Hist. nat. Afr. N.* **29**, 335.
GREGORY, B. D. (1934): *J. Linn. Soc. (Bot.)*, **49**, 531.
HARVEY-GIBSON, R. J. & KNIGHT, M. (1913): *J. Linn. Soc. (Bot.)*, **41**, 305.
HOLLENBERG, G. J. (1940): *Amer. J. Bot.* **27**, 868.
HOWE, M. A. (1916): *Bull. Torrey Bot. Cl.* **43**, 621. — (1918): *Brooklyn bot. Gdn Mem.* **1**, 191. — (1920): *Bull. Torrey Bot. Cl.* **47**, 1.
JANET, C. (1914): *L'alternance sporophyto-gamétophytique de générations chez les algues.* Limoges.
KJELLMANN, R. (1900): *K. svenska Vetensk Akad. Handl.* **33**, 109.
KNIEP, H. (1928): *Die Sexualität der niederen Pflanzen*, pp. 1–544. Jena.
KYLIN, H. (1914): *Svensk bot. Tidskr.* **8**, 33. — (1916): *Ber. dtsch. bot. Ges.* **34**, 257. — (1916*a*): *Z. Bot.* **8**, 97. — (1916*b*): **8**, 545. — (1917): *Ber. dtsch. bot. Ges.* **35**, 155. — (1917*a*): *Naturwissenschaften*, **5**, 84. — (1924): *Acta Univ. lund.* **20**, no. 6, 96. — (1930): **26**, no. 6, 1. — (1937): In *Handbuch der Pflanzenanatomie*, **6**, 1–347.
LEWIS, I. F. (1909): *Ann. Bot., Lond.*, **23**, 639.
MARTIN, M. F. (1939): *J. Linn. Soc. (Bot.)*, **52**, 115.
MATHAIS, W. T. (1928): *Publ. Hartley bot. Lab. Lpool Univ.* **5**, 1.
NEWTON, L. (1931): *A Handbook of the British Seaweeds*, pp. 1–478. London.
OLTMANNS, F. (1898): *Bot. Ztg*, **56**, 99. — (1904–5): *Morphologie und Biologie der Algen.* 1st ed. Jena. — (1922–3): 2nd ed. Jena.
OSTERHOUT, W. J. V. (1900): *Flora, Jena.* **87**, 109.
PHILLIPS, R. W. (1925): *New Phytol.* **24**, 241.
ROSENVINGE, L. K. (1917): *K. danske Vidensk. Selsk. Skr.* 7. Raekke, Nat. og Math. 7, **2**, 155. — (1923–4): **3**, 287. — (1929): *K. danske Vidensk. Selsk.*, Biol. Medd. **8**, 1. — (1931): **10**, 1.
SCHUSSNIG, B. & ODLE, L. (1927): *Arch. Protistenk.* **58**, 220.
SEGAWA, S. (1936): *Inst. Al. Res. Fac. Sc. Hokkaido Imp. Univ.* **1**, no. 2, 175.
SIRODOT, S. (1876): *Ann. Sci. nat.* (*a*), **3**, 146.
SKUJA, H. (1933): *Arch. Protistenk.* **80**, 357. — (1934): *Bot Zbl.* **52**, 188.
SMITH, G. M. (1938): *Cryptogamic Botany*, **1**, Algae and Fungi, pp. 1–545. New York and London.
SMITH, G. M. & HOLLENBERG, G. J. (1943): *Amer. J. Bot.* **30**, 221.
SVEDELIUS, N. (1911): *Svensk bot. Tidskr.* **5**, 260. — (1912): **6**, 239. — (1914): **8**, 1. — (1914*a*): *Ber. dtsch. bot. Ges.* **32**, 1, 49. — (1914*b*): **32**, 2, 106. — (1915): *Nova Acta Soc. Sci. upsal.* Ser. 4, **4**, 1. — (1927): *Bot. Gaz.* **83**, 362. — (1931): *Bot. Zbl.* **48**, 38. — (1933): *Nova Acta Soc. Sci. upsal.* Ser. 4, **9**, 1. — (1937): *Symb. bot. upsaliens.* **2**, 1.
TAYLOR, W. R. (1937): *Marine Algae of the Northeastern Coast of North America*, pp. 1–427. Ann. Arbor.
WESTBROOK, M. A. (1928): *Ann. Bot. Lond.*, **42**, 149. — (1930): **44**, 1012. — (1930*a*): *J. Bot. Lond.*, **68**, 193. — (1930*b*): **68**, 353. — (1935): *Bot. Zbl.* **53**, 564.
WOLFE, J. J. (1904): *Ann. Bot., Lond.*, **18**, 607.
YAMANOUCHI, S. (1906): *Bot. Gaz.* **40**, 401. — (1921): *Bot. Gaz.* **72**, 90.

PIGMENTS IN THE COELENTERATA

By D. L. FOX, Scripps Institution of Oceanography, La Jolla, California, and C. F. A. PANTIN, Zoological Laboratory, Cambridge

(*Received* 20 *April* 1944)

Considering the brightness and variety of the colours of marine invertebrates it is surprising that we still know so little of their biochemical origin and of their biological importance to the animals that possess them. Of no group is this so evident as of the Coelenterata, and the occurrence of large amounts of different pigments in their thin tissues raises the problems of the origin and significance of colour in the most forcible manner. The object of this article is to gather together the limited information we have about these pigments and to draw attention to some of the interesting problems which they present.

Some of the references to the scattered literature are found in the books of Newbigin (1898) and Verne (1926, 1930), and in a recent review by Lederer (1940). A paper by Payne (1931) contains some rather serious inaccuracies with reference to some of the early work. Confusion has sometimes arisen in the literature as to the identity and systematic relations of the animals studied; and as our studies (Fox & Pantin, 1941) of the pigments of the plumose anemone, *Metridium senile*, impressed on us the importance of correct identification not merely of the species but also of the variety investigated, we have wherever possible named species in accordance with Stephenson's monograph on British sea anemones (1928, 1935) and Mayer's *Medusae of the World* (1910). Alternative names used by authors in the works referred to are placed in brackets.

I. MISCELLANEOUS PIGMENTS

Some of the commonest colours in coelenterates are due to pigments whose nature is still uncertain. Among these are the deep blue and the red-brown pigments responsible for the vivid colour of many pelagic coelenterates, such as the Scyphozoa and the Siphonophora, particularly in such predatory forms as *Cyanea* and *Physalia*. In this class we know most about the pigment which is present as minute granules in the scyphozoan *Rhizostoma cuvieri* (Krukenberg, 1882*b*; Blanchard, 1882; Colasanti, 1888). Krukenberg's work was confirmed and greatly extended by von Zeynek (1913). The pigment has three absorption bands, in the red, yellow and green. It is soluble in water, but it is not very stable. It turns violet with alkalis and red in acids, but the latter is an irreversible change. The pigment is altered to a red substance at 50° C. and it is destroyed altogether on boiling. A number of organic solvents such as alcohol, acetone, chloroform, etc., throw it out of solution as a red precipitate. Von Zeynek purified it by fractional precipitation with ammonium sulphate and obtained the following evidence of its composition: C, 51·69 %; H, 7·14 %; N, 10·38 %; S, 1·26 %; ash, 4·0 %. The reactions of the pigment indicate that it is a protein. Krukenberg called it 'cyanein', while Colasanti and von Zeynek refer to it as 'zoocyanin'.

Pigments of very similar properties occur in other Scyphozoa, such as *Aurelia* sp., *Cyanea* sp. and *Cassiopeia borbonica* (M'Kendrick, 1881; Colasanti, 1888; MacMunn, 1890). Colasanti did not, as the title of his papers leads one to suppose, investigate any hydromedusae. He did, however, find a pigment apparently identical with that of *Rhizostoma* in the siphonophores *Velella spirans*, *Porpita mediterranea* and *Physalia pelagica*. Some further observations on the pigment of *Velella* have been made by Haurowitz & Waelsch (1926) and by Kropp (1931). The spectra of all these pigments seem to vary in the distinctness of the absorption bands. Haurowitz & Waelsch noted spectral absorption of the blue *Velella* pigment in the red and violet regions, but discerned no distinct bands. They also state that the blue colour gives way to red in mineral acids, violet in alkali, and turns reddish brown on drying, heating, or treatment with alcohol or toluene. Kropp's observations on the blue pigment in both *Velella spirans* and *Fiona marina*, the nudibranch which devours this siphonophore, agreed in many ways with the findings of Haurowitz & Waelsch. They noted, however, that suspensions of the pigment from either source in distilled water showed a region of absorption in the red, between 655 and 685 mμ, without sharp bands, and exhibited reddish opalescence by reflected light. The pigment was chemically unstable, yellow at neutrality, pink to red in acid, and yielded, on treatment with alcohol, a white flocculent precipitate and a red solution. Observations by Dr J. H. C. Smith (personal communication) on a similar blue coelenterate yielded results like those obtained by Haurowitz & Waelsch and by Kropp, in that treatment with alcohol produced a change of the blue colour to orange or red.

Such blue pigments are classified herein among 'miscellaneous pigments' pending further investigation. We agree with Verne (1930) that they are all probably carotenoid-protein conjugation compounds, similar to those encountered in numerous blue,

green or purple crustaceans (Kuhn & Lederer, 1933; Kuhn & Sörensen, 1938), the blue or purple sea-star *Pisaster giganteus* (Fox & Scheer, 1941), the grey 'sand-crab' *Emerita analoga*, the spines of the soft star *Astrometus sertulifera*, and the skin of the purple nudibranch *Flabellina iodina* (unpublished experiments, D. L. Fox).

We know less about the red-brown pigments. Many Scyphozoa are decorated with these or other coloured pigments instead of blue (Mayer, 1910). Sometimes as in *Cyanea capillata* the colour is blue or brown, according to the variety. Chemically these brown pigments seem to resemble blue ones, and investigation may also show that they are carotenoid-protein conjugation compounds. M'Kendrick (1881) and MacMunn (1890) obtained a brown pigment from the scyphozoan *Chrysaora hysocella*. In its granular condition in the animal and in some of its properties it resembles the blue pigment of *Rhizostoma*. In a brief note Griffiths & Platt (1895) describe a violet pigment which they term 'pelagein' extracted with the lipoids from a scyphozoan, *Pelagia* sp. From their account of its solubility in organic solvents it resembles a carotenoid, but it was said to possess no characteristic absorption bands. They assigned to it an empirical formula containing nitrogen without giving any details of their chemical procedure. Whether or not these red-brown and blue pigments are truly related and are shown to be carotenoid proteins, there seems no doubt that a number of pelagic coelenterates have independently developed similar colouring matters. Pigments of this class may also occur in other phyla. M'Kendrick (1881) suggested that owing to its similar properties the blue colour of the ciliate *Stentor caeruleus* (Lankester, 1873) is related to the blue pigments of the pelagic coelenterates. It is of interest that Emerson (1929) records a red pigment from another ciliate, *Blepharisma* sp., with a three-banded absorption spectrum recalling that of the *Stentor* pigment, but with the bands situated in the short-wave end of the spectrum.

Further work may show that the substances we have discussed above belong to well-known classes. This does not seem to be the case with a well-defined series of pigments found in the skeletons of the Alcyonaria. Krukenberg (1886) originally classed them with the lipochromes as 'rhodophans', perhaps combined with calcium. But most of the latter have proved to be carotenoids, and the alcyonarian skeletal colours differ from these in their properties. In the first place they are much more stable. The lasting nature of the red pigments of the precious coral *Corallium rubrum*, of *Tubipora musica*, and of the reds, oranges and yellows of many other alcyonarians is well known. This is true of the dried corals and of museum specimens preserved in alcohol. The uniquely blue pigment of *Heliopora caerulea* is more intense in the centre of the corallum than on the exposed surface. This may be due to fading under the influence of light. Bourne (1895) found that the grass-green solution of the pigment in dilute ammonia fades to yellow on prolonged exposure to strong sunlight.

The skeleton of the Alcyonaria differs from that of the great majority of the Madreporaria in that the skeleton itself is usually coloured. The pigments are actually present within the calcareous spicules themselves, which in the Alcyonaria are pervaded by a matrix containing a considerable amount of organic substance. They differ entirely from those of the polyp tissues or of the zooxanthellae which these sometimes contain. The pigments from the tissues are carotenoid in nature and are readily soluble in various organic and other solvents, despite Krukenberg's (1887) difficulty in detecting lipochromes in corals.

Alcyonarian coloration is often due to the spicular colours, but occasionally it results from other pigments such as the carotenoids. The red colour of the common European gorgonian *Eunicella verrucosa* is due to the presence of carotenoid droplets in cells of the coenenchyme. The colour rapidly fades on drying and is slowly extracted by alcohol. The spicules of this species are colourless (Studer, 1914). We shall see that this pigment system differs entirely from that of *Corallium*. On the other hand, the eggs and larvae of both the red and white varieties of *Eunicella* have, according to Studer, a very stable pigment which is 'coral red' and not orange-red: perhaps this resembles the spicular pigments of other Alcyonaria.

It is curious that almost all the work on alcyonarian skeletal colours has been done on the blue pigment of *Heliopora caerulea*, an organism which stands apart from the rest of the group. Doubtless the unusual blue coloration of the coral was responsible for the attention it received. The first serious work on the subject was that of Moseley (1876), followed by that of Bourne (1895) and the detailed study of Liversidge (1898). Their accounts are in general agreement. The pigment withstands hot potash and hot 25 % sulphuric acid. It is insoluble in ether, carbon disulphide and cold neutral ethyl alcohol. It dissolves in hot ethyl alcohol or in alcohol which is slightly acid or alkaline. It is readily soluble in formic and glacial acetic acids and to a decreasing extent in higher fatty acids. It is soluble in ammonia and potash. It is insoluble in aqueous solutions of sodium chloride and sodium nitrate, and in hydrochloric acid. Decalcification of the coral by dilute hydrochloric acid frees the pigment in suspension. The pigment is bleached by chlorine and also by reducing agents such as nascent hydrogen. When dried and ignited it leaves a small amount of colourless ash. It is evidently not a carotenoid.

Of the red and yellow colours of the alcyonarian skeleton we know much less. Long ago Vogel (1814) suggested that the colour of the precious

coral *Corallium rubrum* was due to the presence of one part in 88 of ferric oxide. The pigments are in fact organic substances. Nevertheless, the presence of iron is probably significant, as recent work by Durivault (1937) and Ranson & Durivault (1937) showed. Unfortunately, their notes are as brief as they are interesting, and one looks forward to further information in more favourable times.

Durivault found the yellow, garnet-red, red-brown or violet-red pigments of *Alcyonium palmatum*, dredged from the Mediterranean, to be concentrated in the spicules, apparently associated with calcium. Various tests for carotenoids gave completely negative results, and the skeletal pigment failed to dissolve in standard carotenoid solvents. Histochemical tests for carotenoids on decalcified skeletal sections, i.e. with iodine, sulphuric acid and chromic acid, were also negative. The pigment dissolved to some extent in cold potash, and was destroyed by treatment with concentrated hydrochloric, nitric, sulphuric acids or with glacial acetic. When attacked by dilute acids iron could be shown to be released from the dissolving spicules. The iron could be detected by the Prussian blue test and by the thiocyanate reaction. Durivault believes that iron, from the external medium, combines with a calcium organic complex incorporated in the skeletal matrix, and is responsible for the conspicuous colour of this coelenterate. Similar positive tests were obtained on some gorgonians.

Ranson & Durivault reinvestigated the pigment of *Heliopora caerulea* and studied in addition that of *Corallium rubrum* and *Tubipora musica*, both bright red species. In *Heliopora* skeletons no positive tests for carotenoids could be obtained, nor did carotenoid solvents dissolve the blue pigments. The same results were obtained with *Corallium* and *Tubipora*. The skeletons of all three genera yielded positive reactions for iron, and Liversidge had earlier reported phosphate, calcium, magnesium and iron in the ash of *Heliopora*. Ranson & Durivault believe that iron is combined with a calcium organic complex in the skeleton of *Heliopora* just as in *Alcyonium*. They believe that the horny matrix of the spicular skeleton has a special affinity for iron, which combines with the calcium organic complex to give the various pigments characteristic of the group. It will be interesting to learn what kind of pigments these are. In some ways they recall the 'lakes' formed between metals and dyes such as alizarin.

The Alcyonaria are not alone in having 'fast' colours laid down in a calcareous skeleton. A very few madreporarian corals have coloured skeletons, as in some Fungias and certain members of the Eupsammidae (Moseley, 1876), and in *Dendrophyllia nigrescens* (Silliman, appendix to Dana, 1846). The Stylasterina among the hydrocorallines possess many species with fast red or pink skeletons, while in the genus *Distichophora* there are violet, red, orange and brown varieties (Hickson, 1924). We may note also the brilliant and unfading red of the calcareous skeletons of the sessile Foraminifera of the genus *Polytrema*.

Accurate determinations of the iron content of the calcareous skeleton of alcyonarians would be of very great interest. Not only would they throw light on the nature of the pigment but they might provide us with a new line of evidence on the vexed question of whether the palaeozoic tabulate corals are related to the Alcyonaria or not. Clarke & Wheeler (1922) have given a valuable series of analyses of alcyonarian and madreporarian corals, of which a selection are presented below; ferric and aluminium oxides were weighed and determined together in their work. The results are expressed as percentage of total dry weight:

	$(Al+Fe)_2O_3$ %	Organic matter %
Alcyonium carneum	Trace	40·90
Corallium elatior	0·15	3·14
Tubipora purpurea	0·57	2·27
Heliopora caerulea	0·07	1·55
Balanophyllia floridana	0·69	7·35
Acropora cervicornis	0·06	4·32
Oculina diffusa	0·05	2·01
Agaricia purpurea	0·00	1·98

Unfortunately for our present purpose the analyses were apparently made of the whole organism including the soft parts. As might be expected those analyses which included much tissue tend to contain more iron, in both kinds of coral. But in *Tubipora* there is evidently much more iron than can be accounted for in the living tissue as opposed to the skeleton. On the other hand, the mere trace of iron in the *Alcyonium* is interesting in view of Durivault's easily repeatable demonstration of its presence in the spicules of a related species. It should be remembered, however, that iron compounds often have high tinctorial power even when low in iron content, as in the ferrocyanides and the haematins. We must await further analyses.

There is evidence that another class of pigments, the chromolipoids, occurs in the Hydrozoa. Teissier & Volkonsky (1929) studied the calyptoblast hydroids *Serturella gaudichaudi*, *S. gayi*, *Lafoea dumosa*, and *Aglaophenia tubulifera*, and compared the clear green pigments in these species with the yellow, red and blue pigments of the sponge *Clathrina corinacea*. Though the evidence they present is scanty, we believe that it suggests that the *Clathrina* pigments may be related to carotenoids, and that they are not chromolipoids as these authors suppose. The hydroid pigments, on the other hand, were probably true chromolipoid material, as they were insoluble in the usual neutral solvents, took up

vital stains and reduced osmic acid with gradual darkening. The same green pigments were insoluble in water or dilute acids and alkalis, but slightly so in glacial acetic acid. Alkali turned them red-brown, acid dark green. Treatment with formol or sodium bisulphate, or autolysis following death of the colony, cause the pigment to turn brown.

There remain to be mentioned a number of imperfectly known pigments. The yellow 'uranidines' described by Krukenberg (1887) from a number of corals are little more than a name. Sometimes the pigments described do not really belong to the animal from which they were obtained. Symbiotic algae are common among coelenterates, particularly among the Actinozoa. Some of the recorded pigments originate in these plants rather than in the animal host (MacMunn, 1887), as in the 'Antheagrün' of Krukenberg (1881) obtained from *Anemonia sulcata* (=*Anthea cereus*).

Moseley (1873) found a pigment which he called 'actiniochrome' in the red tips of the tentacles of *Tealia felina* (=*Bunodes crassicornis*), though not in the red body of the animal. It was insoluble in all the solvents he used. He did not find it in *Actinia mesembryanthemum* or in *A. rosea* (these are both *Actinia equina*). MacMunn (1885) confirmed this and showed that the same pigment occurred in the tentacles of several other anemones, notably *Anthopleura ballii* (=*Bunodes ballii*) and the violet tentacle tips of *Anemonia sulcata* (=*Anthea cereus*). He found that it was slightly soluble in glycerine. Fulton (1922) concluded that the same pigment occurs in *Actinia bermudensis* and *Condylactis passiflora*. In this he was wrong, for he based his conclusion on similarity of chemical behaviour to a pigment clearly described by MacMunn as actiniohaematin and not as actiniochrome. MacMunn repeatedly differentiated actiniochrome from actiniohaematin: the former never yielded haemochromogen or other haematin derivatives.

II. PYRROLIC PIGMENTS

1. *Haematins*

Whereas the pigments already mentioned still await proper description, we owe to the comprehensive work of MacMunn (1885, 1886) the knowledge that many actiniarians possess haematins and their derivatives. The occurrence of these well-known substances in such simple animals made a great impression upon him. He showed that in many colour variants of *Actinia equina* (=*Actinia mesembryanthemum*), in *Tealia felina* (=*Bunodes crassicornis*), and in other species, there existed a pigment which he called actiniohaematin. Even the white variety of *Metridium senile* (=*Sagartia dianthus*) showed traces of this substance. This pigment yielded a haemochromogen and a haematoporphyrin indistinguishable from those of haemoglobin. Because of this, and because of the capacity of one of the decomposition products of actiniohaematin to undergo reversible oxidation and reduction, MacMunn concluded that the new pigment possesses a respiratory function.

It may be significant that in *Anemonia sulcata*, which possesses symbiotic algae, Elmhirst & Sharpe (1920) detected no haematins, though MacMunn considered there might be a trace of them in the tentacles (as *Anthea cereus*). Elmhirst & Sharpe (1923) confirmed the presence of haematin in *Tealia felina* (=*crassicornis*), and found the concentration of both this and of lipochromes higher in shore specimens than in those from deep water. MacMunn (1890) found that the red variety of the corallinomorph anemone *Corynactis viridis* contains a pigment related to actiniohaematin. Its spectrum shows it to be distinct, but nevertheless it yields products resembling haemochromogen and 'reduced haematin'. In the green variety the spectrum is dominated by another pigment with an absorption band between 507 and 481 mμ in the living tissue.

Roche (1932, 1936 *a*, *b*) brings strong evidence that 'actiniohaematin' is not in fact a pure substance but a mixture derived from the intracellular respiratory enzyme systems such as occur in all animals. He relates it to the cytochromes and to Keilin's 'free intracellular haematins'. He shows that the spectrum corresponds to a mixture mainly of cytochrome *b* and of parahaematin similar to Keilin's protohaematin. It is much stronger in the muscles than in other tissues and it is best developed in those species of actiniarians which have the best developed musculature, such as *Tealia felina*, *Hormathia* (=*Bunodes*) *coronata*, and *Cereus pedunculatus* (=*Heliactis bellis*), while it is more feebly developed in *Anemonia sulcata*, *Actinia equina*, *Adamsia palliata* and in the ceriantharian *Cerianthus membranaceus*, which have a weaker musculature. A search for haematins in the powerful and continually active muscles of the Scyphozoa does not seem to have been made. They are well suited for investigation. In the phylum as a whole it is to be expected that the haematins will prove to be of direct physiological importance to the animal; but they contribute little to visible coloration.

MacMunn (1886) demonstrated the presence of porphyrins in numerous invertebrates, including corals such as *Flabellum variabile* and *Fungia symmetrica*. Moseley obtained 'polyperythrin' (1877) from a number of madreporarian corals, from species of *Actinia* and *Discosoma*, and from *Cassiopeia* and *Rhizostoma* among the Scyphozoa. This substance was shown by MacMunn (1886), in Moseley's own specimens, to be a haematoporphyrin.

2. *Bile pigments*

In addition to a haematin MacMunn (1885) found a pigment which he identified as biliverdin in *Actinia equina* (=*mesembryanthemum*). This pigment is situated beneath the ectoderm and in the

base of the foot, and should not be confused with the green ectodermal coloration of some varieties of this anemone; the latter has since been shown to be a carotenoid protein and will be referred to later. MacMunn also found biliverdin in the green parts of *Tealia felina* (=*Bunodes crassicornis*). The pigment of *Calliactis* (=*Sagartia*) *parasitica* apparently belongs to this group. This is one of the few cases at present known where the main coloration of a coelenterate is probably not due to pigments related to carotenoids, though carotenoids play a minor part in the colour pattern even here (Abeloos & Teissier, 1926). The colour of *Calliactis* varies in individuals from light brown to dark red or even violet. This is chiefly due to the presence in the column of two coloured substances. MacMunn (1885) describes them as brown and purple-red. Subsequently Abeloos & Teissier (1926) described them as red granules found in the superficial tissue, with violet granules in the deeper layers. MacMunn considered that these two substances corresponded to a peculiar new pigment he extracted from the ectoderm, the properties of which he described, and to actiniohaematin, the presence of which he demonstrated in extracts of the endoderm alone. But it is difficult to see how the haematins of the pale endodermal tissues of *Calliactis* can be responsible for any of the granular coloration, and there appears to be another explanation for the two colours.

Abeloos & Teissier found that both red and violet granules were due to the same pigment in different forms. Lederer, Teissier & Huttrer (1940) extended their studies, and have isolated this interesting substance, which they have termed 'calliactine'. Their description of its properties agrees in essentials with that given by MacMunn for his new pigment. The authors found it to be a natural reversible *p*H indicator, acids causing it to assume a yellow-brown colour, passing to golden, then orange, while in alkalis it changes to blue, violet and finally red. It is insoluble in ether or chloroform, but soluble in alcohols and in aqueous systems. In 1 % ammonia its absorption spectrum shows a single, rather symmetrical maximum at about 550 mμ, and in 0·5 % hydrochloric acid a curve of similar shape with maximal absorption in the violet at 450 mμ.

Lederer and his colleagues treated alcoholic solutions of the pigment, slightly acidified with acetic acid, with just sufficient dilute ammonia or potash to turn the colour to blue, whereupon the calliactine was precipitated in the form of rosettes of blue-violet needles. This preparation did not melt until the temperature was raised to 300° C. Its empirical formula is given as $C_{22}H_{22}O_5N_4$ or $C_{21}H_{20}O_5N_4$. It shows an interesting series of colours with minor changes in *p*H, and exhibits, in dilute acid, a yellow-orange fluorescence which disappears in alkali. Numerous physical and chemical reactions lead the writers to conclude that calliactine and its derivatives are related to the bile pigments.

But, in addition to this, MacMunn's original experiments brought to light a most interesting property of this substance. It is easily and reversibly oxidized and reduced. An ammoniacal solution is purple when oxidized and yellow-red when reduced. Contact with air suffices to oxidize it when reduced in solution with ammonium sulphide. Whether the differences of the pigment in the tissues are to be ascribed to differences of *p*H, or to different states of oxidation, or to some other causes is not known. But the possibility that we have here a system of functional importance to the animal is clear. The only pigment which recalls calliactine is the purple substance described by Krukenberg (1882*b*) from *Cerianthus membranaceus*. It is insoluble in lipochrome solvents and weak acids, but soluble in ammonium hydroxide. But MacMunn points out that its spectrum is somewhat different, and whether it will prove to be related to the bile pigments remains to be seen.

III. PURINES

We have found no reference to the existence among coelenterates of complex purines such as the pterines; although Chrometzka (1937) points to the wide distribution of these substances in nature. The simplest purines are not true pigments since they are white in pure crystalline form. But in some Actinozoa a brilliant white in some tissues in contrast with colour in other tissues contributes to the colour effect, as in *Sagartia elegans* var. *venusta*. The substance causing this whiteness is generally unknown, but in *Metridium senile* it has been shown to be due to crystalline purines in the tissues.

Mouchet (1929 *a*, *b*, *c*) claimed that the microcrystalline aggregates in the mesenteric filaments of various anemones, including *Metridium marginatum* and *Actinia equina*, were composed of xanthine. He said they were never uric acid, though he based his conclusions only on the relative solubility of the crystals in dilute acids and alkalis, in ammonia and in piperazine hydroxide. Working on *Metridium senile* Fox & Pantin (1941), however, found uric acid, not xanthine, in amounts of 0·15 % of the dry weight. We showed that negligible quantities of guanine were present, and that no measurable amounts of other purines such as adenine, xanthine or hypoxanthine were demonstrable. We concluded that the microcrystalline deposits in tentacles and mesenteric filaments were uric acid, and not xanthine, the occurrence of which Mouchet believed to be general for anemones. In addition to this the yellow sheets of mucous material excreted, even during starvation, from the column and the mouth of *M. senile* were found to contain no carotenoid material, but to yield considerable quantities of uric acid on treatment with uricase and by other methods (Fox & Pantin). Since Mouchet worked upon European species it is probable that his *M. mar-*

ginatum is identical with our *M. senile*. We therefore disagree with Mouchet, but consider it would be dangerous to argue from *M. senile* to other species. In a note Haurowitz & Waelsch (1926) say they found no uric acid in *Actinia equina*, but give no details.

The large white mass of tissue constituting the 'kidney' of siphonophores such as *Velella* is said to contain needles and granules of guanine (Delaunay, 1931). The analyses of Haurowitz & Waelsch (1926) of the tissues of *V. spirans* show the presence of xanthine, hypoxanthine and possibly guanine also.

IV. MELANINS

Despite their wide occurrence in the animal kingdom the only record of true melanins among the coelenterates prior to the studies of Fox & Pantin is a tentative reference by Abeloos & Teissier to the brown pigment sometimes found in the acontia of *Calliactis parasitica*, and to the slight orange colour sometimes seen in the tentacles of this animal. Verne (1926) and Lederer (1940) cite the older work of Krukenberg (1887), who described a group of 'uranidines' encountered in certain corals, sponges and other invertebrates which turn from yellow to black on death of the animal. Verne has suggested that these pigments may be precursors of melanin.

Our studies on the colours of *Metridium senile* have shown that variants containing brown, grey or black tones in tentacles, stomodaeum, ectoderm of scapus, or endodermal tissues contain brown diffuse, or black granular, melanin in various quantities and combinations. We found further that even in red or white variants a complete tyrosine-tyrosinase system is present, and gives rise to melanin on allowing triturated tissues to stand for short intervals. The melanin formation in such ground tissue systems progresses steadily on continued exposure to atmospheric oxygen, through brown chemically reversible stages to the final black, irreversible condition.

V. CAROTENOIDS

As relatively late as 1922 Palmer said 'there is no definite evidence that coelenterates (sea-anemones, corals, jelly-fish and related animals) contain carotenoids, notwithstanding the brilliant colorations which they exhibit'. He gives one reference each to the early studies of Merejkowski & MacMunn on these animals. Though he verges on overstatement when one bears in mind such observations as those contained in Schulze's (1917) classical monograph on *Hydra*, one must agree that the emphasis which he places on the necessity for further study of pigments in this phylum is still desirable.

Such advances as have been made show that carotenoids and their derivatives are probably the main source of colour in coelenterates. MacMunn (1890) pointed out the wide distribution and similarity of the 'lipochromes' in both the plant and animal kingdom. The most important pigments placed in this ill-defined group are the carotenoids, and probably the majority of the red, orange and yellow colours encountered in the Actinozoa are due to the presence of these substances. Merejkowski (1881, 1883) had already found 'tetronerythrine' and 'zooerythrine' in a number of coelenterates, such as *Actinia equina* (= *mesembryanthemum*), *Aiptasia* sp., *Cereactis* sp. and various Hydrozoa. The pigment of *Aiptasia* might be derived from its symbiotic algae, but in some of the others he was probably dealing with astacene or one of its homologues. Recent studies have shown that many of the actinozoan colours are 'animal' carotenoids, not represented in the plant world, although plant carotenoids have been detected in varying quantities even in species which do not possess symbiotic algae.

There is now evidence of the occurrence of carotenoids in almost every branch of the Coelenterata. Lönnberg (1931, 1933, 1934, 1938; and Lönnberg & Hellström, 1932), in his long series of papers dealing with carotenoids in different animals, includes in the latter a considerable list from this phylum. He characterizes his pigments chiefly by certain of their chemical properties such as solubility in various solvents, colour reactions with antimony trichloride in chloroform, or with strong mineral acids, etc.; in a number of instances he records the positions of absorption bands. Studer (1914) had already brought evidence to show that the orange-red branches of the alcyonarian *Eunicella verrucosa* were coloured by droplets containing carotenoids which were present in the coenenchyme, though absent in the polyps except in local traces. Lönnberg shows beyond doubt the existence of pigments related to carotenoids in the alcyonarians *Alcyonium digitatum*, *Stenogorgia rosea*, *Pennatula phosphorea*, and *Funiculina quadrangularis*. In addition, he finds them in the actiniarians *Tealia* (= *Urticina*) *felina*, *Metridium senile* (= *dianthus*), *Sagartia elegans* (= *undata*), *Actinothoe anguicoma* (= *Sagartia viduata*), *A. lacerata* (= *Sagartia lacerata*), *Bolocera tuediae*, *Stomphia coccinea*, *Hormathia* (= *Chondractinia*) *digitata*, *Halcampa chrysanthellum* (= *duodedimcirrhata*), *Protanthea simplex*; in the madreporarian, *Caryophyllia smithi*; and in the ceriantharian, *Cerianthus lloydi*. He also found them among the Scyphozoa in *Lucernaria quadricornis*. Whether the many colours of the varieties of another stauromedusan, *Haliclystus auricularia*, are entirely due to carotenoids is not known. If so, the variety is astonishing, for there are granular pigments of blue, green, olive, yellow, brown, purple and, rarely, pink colour in the ectoderm (Clark, H. J., 1878). Each individual has but one colour present. Possibly these last pigments are related to the series described above in the pelagic Scyphozoa.

Among the Hydrozoa it has long been known that the red polyp heads of the hydroid *Tubularia indivisa* yield a pigment soluble in alcohol, with an absorption spectrum in the blue between 503 and 468 mμ (MacMunn, 1890). Its solubility in chloroform, ether and alcohol, its absorption spectrum and the colour changes in the presence of sulphuric and nitric acids led MacMunn to decide that the pigment was a lipochrome. Lönnberg & Hellström (1932) have since proved the presence of carotenoids in *Tubularia larynx*. Teissier (1925) also found them in *Clava squamata*. But while they are thus certainly present in gymnoblasts, the evidence for their occurrence in calyptoblast hydroids is conflicting. Lönnberg (1931) points out the necessity of using colonies of these which are free of diatoms. He obtained no evidence of their existence in clean *Antennularia antennina*, and Abeloos & Teissier (1926) make no mention of them in species of *Sertularella*, *Aglaophenia*, and *Lafoea*. In contrast with this Payne (1931) states that she found carotenoids in *Antennularia antennina* and *A. ramosa*, and also in species of *Aglaophenia* and *Halecium*. On the other hand, she found no carotenoids, but only flavones, in five members of the Sertulariidae. Calyptoblasts are not strongly coloured, and possibly their carotenoid content varies with the nature of the planktonic food.

The pigments of the species of *Hydra* are of considerable interest. There exist not merely the commonly recognized green and brown forms, but a whole range of varieties, green, colourless, orange to red, brown or black (cf. Lankester, 1882). The nature of these has been fully discussed by Schulze (1917). In *Chlorhydra viridissima* (= *Hydra viridis*) and other members of the genus the green colour is due to the well-known occurrence of a symbiotic *Chlorella*. Some species of the genus *Hydra* can be temporarily infected with symbiotic algae of various kinds, but the effect does not endure, while in *Pelmatohydra oligactis* even temporary infection is impossible (Goetsch, 1924). In these latter species green colour when it occurs is more usually due, or in *Pelmatohydra* is always due, to green food, such as green chironomids.

Basing his work particularly on *Hydra circumcincta* and *Pelmatohydra oligactis*, Schulze showed that where it was not produced by symbionts coloration in hydras was due to the state of nutrition. Starved animals of all species lose their colour. Well-nourished animals become brown or even black. In all cases the pigments are present as granules of various sizes, particularly in the endoderm cells. Whether all the colours are due to carotenoids is not clear from Schulze's work, but this certainly appears to be the case in red varieties. *Hydra circumcincta*, if fed on ostracods or red copepods, assumes an orange-red colour which increases with the richness of the feeding and vice versa. The colour is due to semi-crystalline granules of carotenoid. When fed on red *Daphnia*, which gave no test for carotenoids (one presumes the colour was due to haemoglobin), the hydras became reddish brown and likewise contained no carotenoid.

Schulze argues that the carotenoids are derived directly from the food: the possibility that they are changed into specific 'animal' carotenoids would not seem to be excluded by his experiments. He also shows that on starvation the carotenoid granules are attacked and gradually disappear. Under certain conditions, as in the body of the mother during egg formation, the red pigment may become brown, to become red again later. Schulze suggests that the apparent care taken to store and use carotenoids indicates that these substances play a special role in metabolism: a remarkable statement considering the time when he made it.

Most of the work already discussed does little more than show that carotenoids are present in the animals, or at most that the chief pigment present belongs to this class. In recent years improved technique has enabled us to study the variety of pigments present in one animal or in different colour forms. The first important work of this kind was that of M. & R. Abeloos-Parize (1926). These authors studied the well-known red, brown and green varieties of *Actinia equina*. They found a red or orange pigment in fine granular form in the ectoderm of the body and tentacles, and a green pigment similarly dispersed in the endodermal tissues. From the brown and green animals they recovered an orange carotenoid, and a red one from the red animals. Subsequently Lederer (1933), and Fabre & Lederer (1934), found in this species a new very unstable carotenoid, actinioerythrin, united as an ester in the red variety. The ester is entirely epiphasic (remaining in petroleum ether when washed with methanol), but becomes entirely and irreversibly hypophasic after alkaline hydrolysis. It is an acidic pigment, destroyed very rapidly on treatment of the alkaline salt (from the hydrolysate) with even such weak acids as acetic or carbonic. The natural ester was readily separated on a Tswett column of calcium carbonate; it melted at 75° C., and gave violet solutions in carbon disulphide, with three absorption bands, at 574, 533 and 495 mμ. In absolute ethanol only one band was reported, at 517–518 mμ. In an impure state, even the ester is very easily bleached by oxidation in the air.

The *green* variant of this anemone possesses, according to these writers, also a red-orange carotenoid displaying the properties of xanthophylls and showing in carbon disulphide two absorption bands, at about 504 and 472 mμ. This new carotenoid exists mostly as an ester, bound to protein, thus assuming a green colour. Both colour variants yielded α- and β-carotenes. The clear blue pigment of the 'marginal tentacles' (acrorhagi?) was concluded not to be a lipochrome.

The investigations of Heilbron, Jackson & Jones

(1935) represent the most careful chemical treatment which has so far been applied to studies of the lipochromes of sea anemones. They investigated four species. They did not specify the colour varieties used, but Dr Jackson kindly informs us in a private communication that in *Metridium senile* it was the red form. The following carotenoids were recovered by chromatography, and in certain instances also by crystallization by these investigators:

Actinia equina. (i) Actinioerythrin. This substance, first described by Lederer, was shown by Heilbron and his colleagues to be an epiphasic ester absorbable on calcium carbonate. It crystallized as red-black needles, m.p. 83–85° C., and had the following absorption bands in carbon disulphide: 574, 538 and 497 mμ. In ethyl alcohol there was one broad band with a maximum at 511 mμ. The presence of this ester indicates that Heilbron was using the red variety of this anemone (cf. Lederer, above). (ii) Violerythrin. This new pigment is an 'acid' or more probably enol-keto carotenoid, *derived by hydrolysis of actinioerythrin,* m.p. 191–192° C. It gave a deep blue colour in benzene, bright blue in pyridine and acetone, violet-red in ether and alcohol and purple in carbon disulphide; in the latter solvent its absorption bands were advanced far towards the red as compared with the parent ester, viz. 625, 576, and 540 mμ. (iii) Xanthophyll (probably a taraxanthin ester) with absorption maxima of 502, 472 and 442 mμ in CS_2. (iv) A trace of carotene.

Anemonia sulcata. (i) Chlorophyll *a* (from symbiotic algae; cf. MacMunn (1885) and Elmhirst & Sharpe (1920)). (ii) Carotene, and (iii) a xanthophyll with absorption maxima at 478 and 450 mμ in carbon disulphide. (Both of these also probably from symbiotic algae.) (iv) Sulcatoxanthin. This is a new xanthophyll insoluble in light petroleum, sparingly so in carbon disulphide, but readily in benzene and alcohol; adsorbed on calcium carbonate or sugar. It has absorption maxima in carbon disulphide at 516, 482 and (450) mμ. It shrinks at 110° C., softens at 125° C., melts at 130° C. and is readily destroyed by mild treatment with alkalis. Its empirical formula is probably $C_{40}H_{52}O_8$.

Metridium senile (=*Actinoloba dianthus*) (red variety). (i) The red pigment is doubtless that described by MacMunn (1890) as having an absorption band between 535 and 565 mμ in the living tissue. Heilbron and his colleagues show that it is an epiphasic ester of low melting temperature, non-crystallizable, which upon saponification at room temperatures yields a red insoluble sodium salt. After hydrolysis the free carotenoid acid crystallized from pyridine in deep violet-red prisms giving a m.p. 195–196° C.; both the free acid and its original ester showed a single absorption maximum in carbon disulphide at 495 mμ. Fox & Pantin (1941) compared this pigment with astacene, from which they showed it to be distinct. They named it 'metridene'. (ii) A small amount of another carotenoid which, in the unhydrolysed state, passed through a calcium hydroxide column in petroleum ether. Its absorption maxima in carbon disulphide lay at 501 and 471 mμ. In addition to these pigments Fox & Pantin found astacene itself and various other carotenoids in this species; they will be mentioned later.

Tealia felina. Heilbron's animals were variable in appearance and in pigment yield, which was small on the whole. There was (i) an epiphasic carotenoid adsorbed as a violet-red zone on calcium carbonate, and resembling actinioerythrin. It was recovered in the form of a violet-black wax, m.p. 65–73° C. and absorption maxima in carbon disulphide; 569, 530, 497 mμ. (ii) An orange-red, low melting epiphasic ester, giving on hydrolysis an insoluble red sodium salt whose free acid separated from aqueous pyridine in black feathery needles, m.p. 205–208° C. Both ester and free acid showed a single absorption maximum at 500 mμ in carbon disulphide.

The pigment situation encountered in *Tealia* by Heilbron and his colleagues is comparable with the findings of Fox & Moe (1938) who investigated the small anemone *Epiactis prolifera* Verrill from the Pacific coast of North America. This species occurs in red, green or brown colour types, and is reported from the Canadian to the Mexican border. The red or orange-red variant was the form investigated. Its extracted carotenoids were entirely epiphasic, consisted of a trace of carotene, possible traces of an esterified xanthophyll and considerable quantities of a red acidic carotenoid ester. The ester and the free acid alike showed a single broad maximum at 500 mμ in carbon disulphide, reminiscent of the corresponding pigment encountered in *Tealia* by the British workers. We do not yet know the underlying factor producing the green colour in the northern variant, but suspect it may be another case of a protein-conjugated carotenoid, since microscopic examination of specimens collected and preserved for some months in formalin by Dr M. W. Johnson, and kindly furnished by him, did not reveal any zooxanthellae. Possession of both the green and red sources of pigmentation may be responsible for the colour of the brown variant.

Valuable as are the careful studies of Heilbron and his colleagues, they still leave the biologist in need of information. We can only surmise from what pigments in the living tissues the carotenoids they examined were derived, and we do not know beyond doubt the colour varieties of some of the anemones used. From the point of view of the biologist wishing to find out the biological significance of the pigments this knowledge is essential. We would urge on all workers in this field the importance of recording the variety of organism from which a pigment is obtained and the source of pigment in the living tissue.

VI. PIGMENTARY BASIS OF COLOUR VARIETIES

If we remain content simply to list the pigments found in an individual coelenterate the main biological problem eludes us. For among these animals we are confronted not only with abundance of pigment but also with an astonishing variety of colours in individuals of the same species. Nor is this variety primarily a question of habitat. In madreporarian corals such as *Porites* and *Pocillopora* (Wood Jones, 1910), in actiniarians (Fox & Pantin), in alcyonarians (Hickson, 1924) and in Stauromedusae

such as *Haliclystus* (Clark, 1878), observers have commented on the way in which numerous colour varieties live side by side under identical conditions. The same is true of pelagic coelenterates, as in the Hydromedusae and the Scyphozoa (Mayer, 1910). Sadly enough this has led to neglect of exact descriptions of pigmentation, for, as Hickson says, colour is considered a 'bad' systematic character. More recently Stephenson has shown the specific importance of range and pattern of coloration.

In some cases variation seems to be of environmental origin. Schulze's work shows that colour variation in *Hydra* depends on nutrition. Elmhirst & Sharpe (1920, 1923), who investigated correlations between the light intensity of the environment and the degree of pigmentation under both natural and artificial conditions, concluded that the anemones *Actinia equina*, *Anemonia sulcata* and *Tealia felina* increase the amount of their pigments, including carotenoids, when in a region of greater light, and vice versa. *Metridium* (=*Actinoloba*) showed no changes of pigmentation under the same conditions. Elmhirst & Sharpe's work showed that in some anemones the pigmentation can vary with external conditions. It is conceivable that in the case of *Anemonia sulcata* their results arose simply from the known variation of the population of symbiotic algae with the light intensity, and hence the change in amounts of algal carotenoids. But this cannot be the case in *Actinia equina* which does not possess these symbionts. Here the work of M. & R. Abeloos-Parize (1926) on the alimentary origin of the carotenoids of the varieties of *Actinia equina* is of great importance. From the brown and the green animals they recovered an orange carotenoid, and a red one from the red animals. Animals raised from the eggs upon a carotenoid-free diet lacked carotenoid pigmentation when grown. Also starved animals regenerated the pharynx and tentacular cycles with only a very feeble carotenoid colouring. Such specimens with carotenoids reduced or absent rapidly recovered the normal pigmentation when fed on a carotenoid-rich diet of shrimps' eggs; furthermore, the regenerating 'browns' and 'greens' reformed their orange pigment, while former 'reds' fed on identical diet recovered their red colour. The work of these investigators indicates two things. First, that in this species the carotenoids are derived ultimately from carotenoids in the diet, and secondly, that there is an important specificity either in the primary selective assimilation of a given pigment, or in the metabolism of a common pigment selected by either variant. Single pure carotenoids added to the diet were not tried.

Like Elmhirst & Sharpe, Fox & Pantin found that the colours of the varieties of *Metridium senile* are much more stable than those of *Actinia equina* seem to be; and though controlled feeding experiments have not been completed, the colour does not appear to change over long periods under a variety of aquarium conditions and of feeding and starvation. Whether this indicates ability in the varieties of *Metridium senile* to construct their carotenoids from simpler substances remains to be seen. But the colour varieties of *Metridium* are certainly stable and maintain themselves side by side in nature apparently on an identical diet, and our analysis showed a complex and surprising set of conditions in them. The varieties prove to be based on a series of biochemical differences of pigmentation which may be classified in four qualitative divisions, viz. (1) white forms, lacking pigment (save for traces in ripe gonads); (2) carotenoid types of pale yellowish, orange, pink or deep orange-red; (3) melanistic types of pale brown, buff, grey, deep brown, or even partly black; and (4) mixtures of (2) and (3). White forms seem to be the most common, with reds a fairly close second, but at some stations the relative order of abundance is reversed. Perhaps to these should be added a bright yellow variety only rarely observed by naturalists on European coasts and not seen by us. Guberlet (1936) reports a yellow form of *M. marginatum* (=*M. senile*?) from deeper water in the north-western coast of the United States.

The substance of a table summarizing the relative abundance and class of carotenoids encountered in various typical colour groups is reproduced below from Fox & Pantin. The chief carotenoids are italicized. For the sake of greater completeness, the relative development of melanin is also recorded. All colour types possess the complete precursor system for melanin formation, the whites and reds resemble cases of 'dominant albinism' (Gortner, 1911).

Colour	Melanins	Carotenoid
White	Little or none	Very little: *Astaxanthin esters* and free astaxanthin
Brown (various shades)	Varying degree	Least. Astaxanthin esters, or metridene esters, carotenes, xanthophylls, xanthophyll esters
Yellow-orange	Little or none	Considerable. *Metridene esters*, *xanthophyll esters*, carotenes, xanthophylls
Red with brown	Varying degrees	Much. *Metridene esters*, or astaxanthin esters
Red	None	Much. *Metridene esters*, sometimes accompanied by free or esterified astaxanthin, free metridene, xanthophylls, carotenes

In the light of the discovery by Kuhn & Sörensen (1938) that astacene is a chemical artifact of the naturally occurring dihydroxy-diketo-β-carotene, we have recorded the presence of *astaxanthin* esters in animals from which we recovered *astacene* in the pigment hydrolysates.

Of the white animals analysed, small quantities of esterified, and still lesser quantities of free astaxan-

thin (i.e. yielding free astacene on treatment with alcoholic potash) were the only carotenoids recognized. Red animals contained from 8 to 10 times as much pigment (in 'astacene equivalents' as measured photometrically) as did the white variety. In the red, furthermore, the carotenoid material was chiefly esterified metridene. β-carotene and astaxanthin esters were usually absent or in minor concentrations, while unesterified metridene (or possibly hypophasic esters), astaxanthin, other relatively heavily oxygenated xanthophylls, and also small amounts of a xanthophyll resembling lutein, were occasionally encountered. Preliminary experiments failed to reveal an oxidase in the white variant capable of destroying the pigment, metridene, from the red form, but other carotenoids were not tested in this way.

Yellow-orange animals showed mixtures of carotenoids, wherein xanthophylls resembling zeaxanthin and lutein, free and esterified, accompanied β-carotene and metridene esters. Esterified zeaxanthin and lutein seemed at least as plentiful as metridene in some such animals. In one large yellow-orange specimen much free zeaxanthin was encountered. Red-brown animals were fairly variable. Some contained esterified and free metridene, accompanied not by free xanthophylls or carotenes but by an esterified xanthophyll with a spectrum resembling that of pectenoxanthin. Another large specimen yielded only small traces of free metridene in the hypophase, possessed no carotenes, ordinary xanthophylls or their esters, but considerable quantities of esterified astaxanthin.

The members of the melanistic (brown or grey) series contained the lowest concentrations of carotenoids (often less than whites), but exhibited the greatest variability in kinds and combinations of these pigments. Some brown, or brown and grey, specimens contained astaxanthin, a fact which was found to be correlated with the possession of pinkish gonads. If the gonads were unripe, astaxanthin esters were lacking, even though other carotenoids might be present. Certain specimens in a sexually unripe condition were found to contain no carotenoids excepting traces of xanthophyll esters, while others might yield small quantities of a carotene, xanthophylls and their esters, and another carotenoid which became acidic on hydrolysis (probably astaxanthin or metridene). Among the carotenoids separated we encountered fractions with absorption spectra agreeing closely with that of β-carotene, other fractions very similar to the lutein or pentaxanthin class, and one which when hydrolysed became hypophasic and possessed a single narrow band in carbon disulphide with maximal absorption at 489–490 mμ.

Our work on these anemones shows how complex the pigment situation can be in each indivudual animal. Several distinct carotenoids may be present in addition to other pigments. Each variant may have its own set of pigments as well as subtle biochemical differences from other varieties.

VII. BIOCHEMICAL SIGNIFICANCE OF THE PIGMENTS

Fox & Pantin (1941) discussed the significance of the pigments in *Metridium*. This has two aspects: the adaptive value of the coloration of the organism in relation to its environment, and the part played by the process concerned with pigment formation in the biochemistry of the organisms. With regard to the first of these we concluded that the colours were not adaptive. This seems to be true not only of many sea anemones but of many marine invertebrates of all kinds (M'Intosh, 1901). The biochemical significance of the pigment-forming processes is still obscure even in the case of the carotenoids, the pigments about which we know most. Merejkowski (1881) considered that the instability of lipochromes in light and in air indicated that they served the animal as respiratory aids. MacMunn (1890) said there was no good evidence for this, and this is still true. Nevertheless, the idea was revived by Studer (1914) to account for the common association which he showed between the white variety of *Eunicella verrucosa* and the presence of symbiotic algae: in the white form he supposed respiration was helped by the algae, in the red form by use of the pigment. But one is tempted to relate the abundance of the algae to the relative absence of light absorption in the white variety. There does appear to be some relation between pigment development and light intensity. Studer himself drew attention to the fact that red varieties of *Hydra* show an increase in carotenoid pigmentation in situations of high light intensity, and we have already mentioned Elmhirst & Sharpe's observations that the pigmentation of some species of anemone is less in dim situations, but some of these possess no algal symbionts so that the effect does not necessarily depend on their presence.

Schulze's observations on *Hydra* and those of the Abeloos-Parizes on *Actinia* show the dependence of the carotenoids on the food supply. It might be supposed that these were simply taken up passively by tissues containing fatty material. We ourselves noted that in pale variants of *Metridium* the small quantities of carotenoids actually present are stored for the most part in the gonads. But the Abeloos-Parizes' work also shows that absorbed carotenoids may undergo change, so that absorption is not entirely a simple process. Further, Schulze found that there was active transport of carotenoid in *Hydra circumcincta* during regeneration and gamete formation. Carotenoid material streams into the maturing egg, where it accumulates as fine granules, not simply dissolved in the yolk. This transport leaves the parent deficient in pigment. This process does not take place in all species; in *Pelmatohydra oligactis* (= *Hydra fusca*) the gonads are not red.

Transference of carotenoid material may be important for development. Teissier (1925) found in the eggs of *Clava squamata* and other Hydrozoa a carotenoid-protein complex which decomposed during early development, releasing free carotenoid. Emerson & Fox (1940) draw attention to the general accumulation of carotenoids in reproductive structures of animals and plants, and to the specific part they have in some cases been shown to play in the fusion of gametes. These authors also found that in the aquatic fungus *Allomyces* the male reproductive cells accumulated carotene, while the female cells remained uncoloured. There is an isolated observation in the Coelenterata which may parallel this. In 1901 M'Intosh stated that the ovaries of *Aurelia flavidula* (= *A. aurita*) are yellow, while the testes are red. Guberlet (1936) makes the same statement, based apparently on her own observations. We have not yet had an opportunity to investigate the chemical side of the sexual differences in pigmentation. The importance of carotenoids for regeneration is shown in Schulze's work in that regeneration in *Hydra* is accompanied by active movement of carotenoid through the tissues. In addition, Okada (1927) demonstrated that the part of the stalk tissue of *Corymorpha tomoensis* which is rich in red pigment is also tissue which has special power of regeneration.

Little as we know of the function of the carotenoids in any individual, we know even less of the significance of varieties containing different pigments. At first glance it is tempting to assume that the presence or absence of pigments such as the carotenoids in the varieties of a coelenterate may depend merely on their respective capacity to store, or to oxidize and destroy these very labile compounds. Thus we might expect a tendency in purely melanistic forms which produce melanin by an oxidative attack on tyrosine, to exclude storage of the easily oxidized carotenoids. In *Metridium* there is in fact some ground for this expectation because, on the whole, in those varieties of *Metridium* which are richest in carotenoids melanin formation is less pronounced. On the other hand, this distinction is not absolute, and some varieties of *Metridium* store both the oxidation product melanin and the carotenoid pigments. But it is factors of this sort which determine the colour of varieties among some plants, the nearest approach to the colour varieties of coelenterates (Scott-Moncrieff, 1937).

Whatever causes pigment to accumulate, the pattern of its deposition is closely related to structural differentiation throughout the Coelenterata. Stephenson (1928) has pointed out the intimate correlation of the complex and beautiful patterns of sea anemones with the underlying structure in them. We may cite as examples the relation of the red pigment on the disk of *Tealia felina* var. *coriacea* to the endocoelic spaces round the base of the tentacles, or the accumulation of endodermal melanin along the mesenteric insertions of some varieties of *Metridium senile*. In different varieties of anemones, structurally differentiated regions may be picked out in this way, sometimes by one colour, sometimes by another. Such a relation between organization and pigmentation is common throughout the coelenterates and many other groups of animals (Tylor, 1886). It is particularly evident in marine organisms such as the brightly coloured Polychaeta. Among these Allen (1927) has shown a peculiar relation between autotomy and pigmentation in *Autolytus macrophthalma*. The worm has a series of natural breaking points. These occur in a regular manner every few segments, and their position is marked by the absence of yellow pigment in the septa.

Pattern of pigmentation is not due solely to the casual accumulation of a particular pigment in a particular kind of cell. This is most evident when we compare colour varieties. Thus, the red variety of *Metridium* usually contains carotenoid throughout the whole column, but Fox & Pantin record examples in which the pigment is restricted to the scapus and again others in which it is restricted to the capitulum. Each particular tissue has its own colour processes, and the same tissue is linked with different processes in different varieties.

Throughout the Coelenterata great variation of pigment pattern is common. Apart from the Actinozoa we may call to mind varied red, yellow and brown patterns among Scyphozoa such as *Cyanea capillata* and *Cassiopeia* spp., or the varieties seen in Hydromedusae such as *Turris pileata* (Mayer, 1910). As Stephenson says of anemones, it is the plan of the pattern which is characteristic of the species rather than the particular coloration.

VIII. GENETIC SIGNIFICANCE OF COLOUR VARIETIES

The origin of such variations of pigment pattern is a genetical question on which no work seems to have been done. In many coelenterates the colour varieties, although connected by intermediates, centre round several distinct types and are thus true cases of polymorphism (Huxley, 1942). The condition in *Metridium* is typical of many. Fox & Pantin imply that since the colour patterns of *Metridium* seem to be non-adaptive they might be due to unselected mutations. The system of coloration in *Metridium* shows a rather striking parallel to the system of the human blood groups. The latter depends on a few simple genetic differences (Wright, 1940). Colour differences in coelenterates recall colour differences in flowers, and like these might be due to simple genetic differences (cf. Scott-Moncrieff, 1937). In both the anemone colour system and the blood group system there is no obvious functional advantage of one variant over another. In both systems several variants are maintained in fairly large proportion, but the proportions vary in isolated communities. Wright discusses this isolation

in connexion with the blood groups. In coelenterates there are many causes which favour it. Not uncommonly in this phylum the effective breeding individual is an asexually produced clone which, barring geological accident, can live on the same site indefinitely. As in *Metridium*, an entire ecological situation may be occupied by the asexual products of a few animals, and the chance of any of the immense number of sexually produced larvae obtaining a foothold in an already occupied site is remote. In the case of the blood groups Wright argues that the system results from random local differentiation. The same argument may apply to the colour systems of organisms like *Metridium*.

However, there is doubt whether unselected mutation will suffice to give polymorphism. Ford (1934) points out that there is evidence that the coexistence of several varieties can apparently only be maintained if this condition offers a selective advantage. Whether the advantage is directly concerned with the possession of colour or with some other indirect effect on the organism, the balance of advantage between a gene and its allelomorphs must be extraordinarily exact if there is to be no selection. Unfavoured by natural selection the rate of diffusion of a gene through a population may be far too slow to account for any observed cases of polymorphism (Ford, 1940). Diffusion would be especially slow in anemone polymorphism because the interval between generations may be so great.

There is no direct evidence as to the nature of any advantage conferred by colour polymorphism in coelenterates. Ford (1934), Fisher (1930) and Huxley (1942) discuss cases of polymorphism which seem to resemble that of *Metridium*; the shell colour of *Cepaea* (=*Helix*) *hortensis* and *nemoralis* (Diver, 1939) is an example. In such cases the polymorphs show dominance to a universal recessive. They appear as heterozygotes, for these have an advantage over the homozygote because the genes concerned are closely linked with lethals, so that the homozygotes are less viable. If the polymorphs of *Metridium* compare with these, we might expect that the common white variety was a universal recessive and not a dominant albino as Fox & Pantin (1941) earlier suggested. Breeding experiments to determine this would be interesting. Whatever the genetical mechanism of such colour polymorphism (cf. Haldane, 1930), it should be remembered that it is very common among invertebrate organisms. One may refer to colour varieties which live side by side in sponges such as *Halichondria panicea*, in Turbellaria such as *Polycelis nigra*, in nemertine worms such as *Lineus ruber*, as well as in various molluscs.

The genetical problem of the colour varieties of coelenterates may be stated as follows. As exemplified in *Metridium* the varieties seem to be based on the intermixture of a few simple pigment systems in different proportions. The resulting varieties live side by side without obvious relative advantage to one or other. The proportions of the varieties vary, but most are well represented on any one ecological site. The animals are peculiar because the prevalence of asexual reproduction often means that the effective breeding individuals are few and of immense longevity; while the sexually produced larvae occur regularly in immense numbers with only a remote hope of settling. Any hypothesis put forward must take account of the widespread character of colour polymorphism among the simpler invertebrates: special hypotheses to meet particular cases are therefore improbable.

IX. SUMMARY

(1) The nature, variety and significance of coelenterate pigments is discussed. They may be grouped under: miscellaneous pigments, the pyrrolic series, purines, melanins, carotenoids. (2) The nature of several well-known pigmentations is not fully understood. Chief of these are the deep blues and browns of pelagic coelenterates, possibly due to carotenoid-protein compounds; and the remarkable stable colours of the spicules of Alcyonaria. The latter appear to be combined with calcium and iron. (3) Of the pyrrolic series haematins are common. They play little part in coloration and are probably simply part of the intracellular respiratory system. Bile pigments also occur and attention is drawn to the interesting pigment of this kind present in *Calliactis parasitica*. (4) Purines only contribute white to the colour pattern of some anemones. They occur as uric acid and possibly as other purines. (5) Black and brown melanins occur in *Metridium senile* and probably in other anemones. Even white anemones may possess a complete tyrosinase system. (6) It seems that coelenterate colours are due chiefly to carotenoids. These are found in species from all parts of the phylum. Though the evidence is incomplete, the carotenoids are in some cases known to be derived from the food (*Hydra*, actinians). Sometimes at least, however, they are altered to 'animal' carotenoids, as in the 'metridene' from *Metridium*. Different colour varieties may form chemically different carotenoids. (7) Detailed study of actinians shows a surprisingly complex pigment situation. An individual may possess several distinct carotenoids, and different varieties of the same species are characterized by different series of pigments. (8) Little is known of the biochemical significance of the pigments. There are indications that the presence of carotenoids is in some cases related to light intensity. There is also evidence that these pigments are of importance in gamete formation and in regeneration. The biochemical significance of the varietal colour differences is not known. Attention is drawn to Stephenson's observations on the importance of range of colour and of colour pattern in relation to morphological axes of differentiation. (9) The genetic significance of the existence of colour varieties is discussed. True polymorphism with respect to colour is common in coelenterates. No genetic explanation can be offered without further experiment. The resemblance of the colour systems of *Metridium* to the blood group systems in man is commented on. Attention is drawn to the widespread character of colour polymorphism in the lower invertebrates.

X. REFERENCES

ABELOOS, M. & TEISSIER, G. (1926): *Bull Soc. zool. Fr.* **51**, 145.

ABELOOS-PARIZE, M. & R. (1926): *C.R. Soc. Biol., Paris*, **94**, 560.

ALLEN, E. J. (1927): *J. Mar. Biol. Ass.* **14**, 869.

BLANCHARD, R. (1882): *Bull. Soc. zool. Fr.* **7**, 402.

BOURNE, G. C. (1895): *Philos. Trans.* **186**, 455.

CHROMETZKA, F. (1937): *Ann. Rev. Biochem.* **6**, 211.

CLARK, H. J. (1878): *Smithson. Contr. Knowl.* **23**, Art. 242, p. 1.

CLARKE, F. W. & WHEELER, W. C. (1922): *Prof. Pap. U.S. Geol. Surv.* no. 124, p. 1.

COLASANTI, J. (1880): *Atti accad. med. Roma*, **12**. — (1888): *Das blaue Pigment der Hydromedusen.* Untersuch. zur Naturlehre des Menschen und der Tiere. Herausgegeben von J. Moleschott, **13**, 471.

DANA, J. D. (1846): *Structure and Classification of Zoophytes.* Philadelphia: Lea and Blanchard.

DELAUNAY, H. (1931): *Biol. Rev.* **6**, 265.

DIVER, C. (1939): *J. Conchol.* **21**, 91.

DURIVAULT, A. (1937): *C.R. Soc. Biol., Paris*, **126**, 787.

ELMHIRST, R. & SHARPE, J. S. (1920): *Biochem. J.* **14**, 48. — (1923): *Ann. Mag. Nat. Hist.* 9th series, **11**, 615.

EMERSON, R. (1929): *J. Gen. Physiol.* **13**, 159.

EMERSON, R. & FOX, D. L. (1940): *Proc. Roy. Soc.* B, **128**, 275.

FABRE, R. & LEDERER, E. (1934): *Bull. Soc. Chim. biol., Paris*, **16**, 105.

FISHER, R. A. (1930): *Amer. Nat.* **64**, 385.

FORD, E. B. (1934): *Mendelism and Evolution.* London: Methuen and Co. — (1940): Polymorphism and Taxonomy, pp. 493–513. *The New Systematics*, ed. by J. Huxley. Oxford: Clarendon Press.

FOX, D. L. & MOE, C. R. (1938): *Proc. Nat. Acad. Sci., Wash.*, **24**, 230.

FOX, D. L. & PANTIN, C. F. A. (1941): *Philos. Trans.* B, **230**, 415.

FOX, D. L. & SCHEER, B. T. (1941): *Biol. Bull. Woods Hole*, **80**, 441.

FULTON, J. F. (1922): *Quart. J. Micr. Sci.* **66**, 339.

GOETSCH, W. (1924): *Z. wiss. Biol.* **1**, 660.

GORTNER, R. A. (1911): *J. Biol. Chem.* **10**, 113.

GRIFFITHS, A. B. & PLATT, C. (1895): *C.R. Acad. Sci., Paris*, **121**, 451.

GUBERLET, M. L. (1936): *Animals of the Seashore.* Portland: Metropolitan Press.

HALDANE, J. B. S. (1930): *Amer. Nat.* **64**, 87.

HAUROWITZ, F. & WAELSCH, H. (1926): *Hoppe-Seyl. Z.* **161**, 300, and note on p. 318.

HEILBRON, I. M., JACKSON, H. & JONES, R. N. (1935): *Biochem. J.* **29**, 1384.

HICKSON, S. J. (1924): *An Introduction to the Study of Corals.* London: Longmans, Green and Co.

HUXLEY, J. (1942): Evolution. *The Modern Synthesis.* London: George Allen and Unwin.

KROPP, B. (1931): *Biol. Bull. Woods Hole*, **60**, 120.

KRUKENBERG, C. FR. W. (1881): *Vergleichend-physiologische Studien*, 1. Reihe, 5. Abt. p. 38, Das Antheagrün. Heidelberg: Carl Winter. — (1882*a*): *Vergleichend-physiologische Studien*, 2. Reihe, 3. Abt. p. 62, Ueber das Cyanëin und das Asterocyanin. Heidelberg: Carl Winter. — (1882*b*): *Vergleichend-physiologische Studien*, 2. Reihe, 3. Abt. p. 72. Beiträge zur Kenntniss der Actinienfarbstoffe. Heidelberg: Carl Winter. — (1887): *Vergleichend-physiologische Studien*, 2. Reihe, 4. Abt. p. 172. Die Farben der lebenden Korallen des Rothen Meeres. Heidelberg: Carl Winter.

KUHN, R. & LEDERER, E. (1933): *Ber. dtsch. chem. Ges.* **66**, 489.

KUHN, R. & SORENSEN, N. A. (1938): *Ber. dtsch. chem. Ges.* **71**, 1879.

LANKESTER, E. RAY (1873): *Quart. J. Micr. Sci.* **13**, 139. — (1882): *Quart. J. Micr. Sci.* **22**, 229.

LEDERER, E. (1933): *C.R. Soc. Biol., Paris*, **113**, 1391. — (1940): *Biol. Rev.* **15**, 273.

LEDERER, E., TEISSIER, G. & HUTTRER, C. (1940): *Bull. Soc. chim. Fr.* **7**, 603.

LIVERSIDGE, A. (1898): *J. Roy. Soc. N.S.W.* **32**, 256. — (1899): *Chem. News*, **80**, 29.

LÖNNBERG, E. (1931): *Ark. Zool.* **22**A (no. 14), 1–49. — (1933): **25** (no. 1), 1–17. — (1934): **26**A (no. 7), 1–36. — (1938): **30**A (no. 6), 1–10.

LÖNNBERG, E. & HELLSTRÖM, H. (1932): *Ark. Zool.* **23**A (no. 15), 1–74.

M'INTOSH, W. C. (1901): *Ann. Mag. Nat. Hist.* 7th series, **7**, 221.

M'KENDRICK, J. (1881): *J. Anat. Physiol.* **15**, 261.

MACMUNN, C. A. (1885): *Philos. Trans.* B, **176**, 641. — (1886): *J. Physiol.* **7**, 240. — (1887): *Quart. J. Micr. Sci.* **27**, 573. — (1890): **30**, 51.

MAYER, A. G. (1910): *Medusae of the World*, **1**, **2** and **3**. *Publ. Carneg. Instn.*, no. 109.

MEREJKOWSKI, C. DE (1881): *C.R. Acad. Sci., Paris*, **93**, 1029. — (1883): *Bull. Soc. zool. Fr.* **8**, 81.

MOSELEY, H. N. (1873): *Quart. J. Micr. Sci.* **13**, 143. — (1876): *Philos. Trans.* **166**, 91. — (1877): **17**, 1–23.

MOUCHET, S. (1929*a*): *C.R. Soc. Biol., Paris*, **102**, 733. — (1929*b*): *Bull. Soc. zool. Fr.* **54**, 345. — (1929*c*): *L'excrétion chez les Actinies.* Stat. Océanographique de Salammbô, Notes, no. 15, pp. 1–15, Salammbô.

NEWBIGIN, M. (1898): *Colour in Nature.* London: John Murray.

OKADA, Y. K. (1927): *Arch. Zool. exp. gén.* **66**, 497.

PALMER, L. S. (1922): *Carotinoids and related pigments.* New York: A.C.S. Monograph Series, Chemical Catalog Company, Inc.

PAYNE, N. M. (1931): *J. Mar. Biol. Ass. U.K.* **17**, 739.

RANSON, G. & DURIVAULT, A. (1937): *C.R. Soc. Biol., Paris*, **126**, 1149.

ROCHE, J. (1932): *C.R. Soc. Biol., Paris*, **111**, 904. — (1936*a*): **121**, 69. — (1936*b*): *Ann. Rev. Biochem.* **5**, 463.

SCHULZE, P. (1917): *Arch. Biontol., Berl.*, **4**, Heft 2, pp. 39–119.

SCOTT-MONCRIEFF, R. (1937): The biochemistry of flower colour variation, pp. 230–43. *Perspectives in Biochemistry*, ed. by Needham and Green. Cambridge: University Press.

STEPHENSON, T. A. (1928, 1935): *The British Sea-Anemones*, **1** and **2**. Roy. Soc. London: Dulau and Co.

STUDER, T. (1914): *Zool. Anz.* **43**, 449.

TEISSIER, G. (1925): *Trav. Stat. biol. Wimëreux*, **9**. Also in *C.R. Congrès Avanc. Sci.* 48e Session, Liége, p. 481.

Teissier, G. & Volkonsky, M. (1929): *Bull. Soc. zool. Fr.* **54**, 599.
Tylor, A. (1886): *Coloration in Animals and Plants.* London: Alabaster, Passmore and Sons.
Verne, J. (1926): *Les pigments dans l'organisme animal.* Paris: G. Doin. — (1930): *Couleurs et pigments des êtres vivants.* Paris: Librairie Armand Colin.
Vogel (1814): *Ann. Chim. (Phys.)*, **89**, 113.
Wood Jones, F. (1910): *Corals and Atolls.* London: Lovell Reeve and Co.
Wright, S. (1940): The statistical consequences of mendelian heredity in relation to speciation, pp. 161–183. *The New Systematics*, ed. by J. Huxley. Oxford: Clarendon Press.
von Zeynek, R. (1913): *Mh. Chem.* **34**, 581–621.

THEORETICAL ASPECTS OF BACTERIAL CHEMOTHERAPY

By HENRY McILWAIN, Leverhulme Research Fellow
Department of Bacterial Chemistry (Medical Research Council), University, Sheffield

(*Received* 22 *June* 1944)

I. INTRODUCTION

The term chemotherapy, in its original meaning and in its orthodox current usage, does not apply to all therapeutic treatment with chemical substances. Such treatment is described as chemotherapy when it is carried out in a particular system which can be defined only in biological terms: namely, that of organisms related as host and parasite. The special problems of chemotherapy are thus differentiated from those of pharmacology, not by their predominantly chemical interest but by their different subject-matter and viewpoint; any complete study of a chemotherapeutic agent includes both its pharmacological and its parasitological investigation. The present account attempts to present chemotherapy in a formal framework which overtly expresses it as a subject concerning the properties and the various interactions of drug, parasite, and host. Though primarily concerned with the newly developed subject of bacterial chemotherapy, the approach is derived from knowledge which existed prior to the successful use of antibacterial chemotherapeuticals and reference is necessarily made to this knowledge. In its arrangement the account departs from current presentations of chemotherapy in which the subject-matter is arranged either in terms of the chemical structures of the agents, or in terms of the diseases against which they are employed. Such arrangements have accorded with neither theoretical nor practical approaches to chemotherapy. Thus it has frequently been found that agents developed against one disease have been of greater value against other, very different, ones; that compounds very different chemically have been effective against the same infection, or similar compounds, against entirely different infections. Such findings have appeared disorderly while chemical or pathological disciplines have been imposed upon chemotherapy, and in the past they have intensified empirical approaches to a subject which is extremely rich in its theoretical associations.

As the factors conditioning pathogenesis are many, so also are the means by which the process is open to disturbance. Though chemotherapy concerns only those cases in which disturbance is by administration of substances to a host, many ways can be envisaged for such interference and more than one type of chemotherapeutic action has been established. Chemotherapy is usually regarded as distinct from those methods which involve pre-treatment of the hosts or parasites, as do disinfection, ecological methods, and active immunization, though obvious borderline cases exist. Knowledge of the factors involved in pathogenesis is incomplete, but certain ones which have been shown to be involved in chemotherapeutic actions may be briefly discussed.

(1) Normal infection is by a relatively small inoculum. This may be of some thousands of organisms (though of virulent bacteria, a single cell can lead to fatal results) but is very much smaller than that necessary to cause, without multiplication of the parasite, the effects of the disease. Growth of the parasite is thus invariably an important factor. (2) The mere presence in the host of an amount of inert substance equivalent to that of the parasite at a time when infection is fatal is not considered in itself to be adequate to explain the effects upon the host; specific activities of the parasite, e.g. in elaborating toxins, are involved in pathogenicity. (3) Most bacteria do not grow well in animals, though, when suitably treated, tissues and tissue fluids make excellent media for growth of many organisms, pathogenic and non-pathogenic. Agents antagonistic to micro-organisms are thus present in potential hosts; they may not merely prevent growth of the parasite, but may kill it. (4) Pathogens are nevertheless capable of growing in normal, untreated, animal tissues. They have thus avoided or counteracted the antibacterial agencies of the host.

Chemotherapy normally results in elimination of the parasite from the host, though this may be incomplete. The above considerations are not exhaustive but show that this may have occurred by killing or stopping growth of the parasite; by affecting its production of toxic substances; or by disturbing mechanisms by which the parasite normally avoids the host's antibacterial agencies. These and other interactions are here presented formally by considering separately the properties of components of the chemotherapeutic system, and their behaviour in simpler combinations.

II. DRUG

The class of substances which are chemotherapeutic agents have not been defined in terms other than chemotherapeutical. Inorganic and organic compounds have proved effective, and of organic compounds, a great structural variety. Biochemically,

the agents acting upon a given parasite may have the complexity of enzymes or the relative simplicity of coenzymes or substrates. A great part of chemotherapeutic endeavour has nevertheless consisted of experiments in which the main variable has been the structure of the potential agent, and attempts to rationalize the results have consisted in comparing those structures, or other physical or chemical properties of the agents, with their effects in the complete chemotherapeutic system. This is understandable, as in the preparation of organic substances their structures are most immediately under experimental control. Working upon chemical analogies, with the minimum of specific biological knowledge, the proportion of successful agents has been extremely small, but to such methods are owed the compounds which initiated chemotherapy and provided material for the theorist.

Work (1940) emphasized that structural criteria for antiplasmodials are few; there are exceptions even to the generalization that such agents are bases of molecular weights between 300 and 400. Dewing, Gray, Platts & Stephenson (1942) observed in the preparation of antibacterials that accumulation in one molecule of two or more groups which individually could confer activity, frequently led to inactivity. The probability of inactivity extends also to the multiplication of a single type of active group in one molecule: polysulphanilamides were found to be relatively inactive (Mann & Watson, 1943). Combination of groups akin to those active in other connexions led to plasmoquin, but only after the trial of 'every conceivable' variation (Schulemann, 1932). When a series of compounds of a given structural type is examined with respect to a particular chemotherapeutic action, optimal activity may be found to be associated with a maximum in certain physical properties, often in more than one such property; Ferguson (1939) has emphasized their frequent and fundamental interconnexion, and incidentally the insecurity of conclusions based upon examination of only one property of this type. The correlations are not very extensive (Rawlings, Sweet & Joslyn, 1943; Page & Robinson, 1943; Hager & Grub, 1942).

The general conclusion may be anticipated that such results are due to the manifold demands and responses of various systems of the various components involved in chemotherapy. After preliminary biological study of specific components, approximations to the chemical desideratum of correlation between structure and action have been reached in a few notable cases (cf. §§ V 1 a_2; VI).

III. PARASITE

Bacteria as a whole possess the following general properties which appear relevant to the parasitism of certain of their members and to the action of antibacterial chemotherapeuticals. (1) They are capable of growth in a great variety of environments and of adaptation to circumstances which are new to them, such as the presence of most drugs. Higher animals are not only constantly in contact with them because of their widespread distribution but also through the many species of bacteria which are commensal with animals; certain micro-organisms formerly regarded as commensal are now known to be in symbiotic relationship with higher animals and the danger of too effective inhibition of micro-organisms by a chemotherapeutic agent is a real one. (2) Most pathogenic bacteria grow rapidly, at rates much greater than those of animal tissues; the detailed course of growth is considered later. Their reproduction is asexual; a single organism can cause infection. They are extremely small, which increases their opportunity for passive distribution but limits their active distribution; it also implies that from the average inoculum relatively enormous growth is necessary before they become offensive, and this gives considerable scope to the action of chemotherapeuticals. (3) Though in basic metabolic processes bacteria show many properties extremely similar to those of the organisms which they parasitize, they show also obvious differences including those which make certain of their members inimical to other forms of life. Even non-parasitic species produce substances toxic to higher animals. Many limitations but significant opportunities are thus presented to achieving the differential action which is necessary in chemotherapy.

It is not very profitable to consider what taxonomic groups among the bacteria contain pathogens, as pathological characters themselves—or physiological ones related to pathogenesis—are frequently used as differentiating characters in bacterial classification. Certain properties of bacteria, other than their behaviour in animals, have, however, been found to discriminate between pathogenic and non-pathogenic strains: heat resistance of the streptococci, and to some extent their fermentative powers; the reactions of pneumococci to inulin and bile; and, especially, the fermentative characters of the colon-typhoid group. Adaptive reasons can be seen for the reactions of many pathogens to animal constituents, but rational connexions in most of the other cases have not been given.

IV. HOST

Chemotherapy has in its relatively short career been concerned mainly with human infections and to a lesser degree with those of domestic animals and carriers of human diseases. Laboratory studies tend to use the smaller, more convenient, animals as experimental hosts; few comparative studies have been made in chemotherapy proper but more are available with respect to the pharmacology of chemotherapeutic agents.

Animals have evolved in association with bacteria and present many characters which can be referred to that association (§ VII); certain more general points may be considered here. The tissues of higher animals are from many points of view very favourable

to bacterial growth; in contrast to other parts of the body, or to their environment in general, a concentration of nutrient substances under controlled conditions is available to the bacteria. These concentrations, however, vary notably in different animal species and with growth of the animal; this is related to chemotherapy, as many substances of the host are antagonistic to chemotherapeutic action (cf. McIlwain, 1943*e*). Animals have general protective devices which are apparently not directed primarily against bacteria but which nevertheless condition the mutual interaction of the two types of organism: the barrier of the skin, and the collecting and disposing of any particles which reach the mouth and lungs, or tissues. The general well-being of an animal can condition the progress of infection and chemotherapy: deficiency in vitamins A, C (cf. Cannon 1940), some of the B group and possibly others can predispose animals to infection. The vitamins themselves thus have, in the specific circumstances of infection in a nutritionally deficient host, many of the characters of chemotherapeutic agents. Their action then represents a type suggested (but not established) as being the mode of action of orthodox chemotherapeuticals: namely, an increase in the host's native antibacterial potencies. Non-specific irritation of the host can stimulate various mechanisms and result incidentally in greater antibacterial activity; substances or conditions causing this also simulate to some degree the effects of chemotherapeuticals (e.g. Andrewes, King & Ende, 1943).

V. DRUG AND PARASITE

1. *Effects on growth of the parasite*

Reasons have been given for considering the growth of the parasite to be an important event in the chemotherapeutic system, whether or not the primary effect of the agent is upon such growth. In addition, many agents have been found to act through effects upon growth (§ VIII). This describes a biological characteristic of their action. In itself the description is not of great assistance to the chemical aspects of chemotherapy, but the phenomena concerned are susceptible to further analysis. Of the various methods of further study, the present account deals particularly with antagonism of the action of the agents, and with the action of the agents upon the course of growth of bacterial cultures.

(a_1) *Antagonism of antibacterial actions of chemotherapeutic agents*

Drug antagonism is an established technique in pharmacological studies and practice, but its extensive use in chemotherapy is more recent (cf. Fildes, 1940*b*; McIlwain, 1942*a*, 1943*a*). The basic type of observation in such approaches to chemotherapeutic action is that an inhibition of bacterial growth or other activities, caused by adding an agent to a culture, may be prevented by the addition of further materials. A common procedure, developed from nutritional studies, is to search for a crude natural material which in relatively small quantities can antagonize the effects of the inhibitor. Such a material may not be found, and the technique is thus not of universal application. When, as in the cases discussed below, antagonism has been found, normal methods of separating nutritional factors from natural sources have usually been applied to the antagonistic material and have resulted in the ascribing of its action to particular substances or groups of substances. Their connexion with the action of the inhibitor is shown primarily by the organism's lesser need of them in the absence of the inhibitor, and often also by the finding that their antagonistic effects are limited to a given inhibitor or class of inhibitors (McIlwain, 1942*e*). The findings are thus primarily in terms of substances, and their interpretation in terms of processes, which is necessary to apply the findings to the understanding of chemotherapeutic action, are indirect or dependent upon other evidence. In particular, their interpretation is more complex and less direct than is the interpretation of straightforward nutritional studies (McIlwain, 1943*b*, *e*; Valko & Dubois, 1944). This is illustrated by the examples of the following paragraphs. In addition, Gaddum (1943) has discussed formally two important types of antagonism: by combination of antagonist with the drug, and by competition between the two compounds.

(α) As *and* Hg *derivatives.* Voegtlin (1925), following a suggestion of Ehrlich (1909), gave evidence for a specific relationship between the action of arsenic compounds and thiol derivatives upon trypanosomes. Voegtlin found that the action of the drug was antagonized by thiol derivatives, and further, on the basis of finding thiol derivatives in trypanosomes, made the major transition of suggesting that the drugs acted by combination with thiol groups of importance to the parasite. Fildes (1940*a*, *b*) developed these views and also wider, metabolic, ones in relation to bacterial inhibition in his interpretation of the action of sulphonamides (see below) and his study of the antibacterial action of mercuric salts. In the latter case, evidence that the inhibitor affected thiol derivatives of importance to the bacterial cell was given by quantitative similarities between the interaction of mercuric salts and thiol derivatives in the absence of bacteria, and their interaction in inhibiting and restoring growth of *Bact. coli* (cf. Valko & Dubois, 1944).

(β) *Sulphonamides.* Stamp (1939), investigating the antistreptococcal action of sulphanilamide, and Green (1940), its action upon *Brucella abortus*, separated from the organisms themselves material which antagonized the action of the drug. Woods (1940), following Stamp's procedure with yeast as a more accessible source, purified the extracts and examined their chemical properties, which were in certain salient points similar to those of the inhibitor. Relatively small amounts of the concentrates were

capable of antagonizing a given amount of sulphanilamide: antagonism was thus unlikely to be by combination. To counteract increasing quantities of sulphanilamide greater amounts of the concentrates —roughly in proportion to the increased quantities of sulphanilamide—were required. Such competitive inhibition of biological processes was already known in enzymic reactions, and it was suggested that the concentrates might be the substrate of an enzyme, whose product was essential to the cell, and whose functioning was inhibited by sulphanilamide by virtue of its similarity to the substrate. This enabled suggestions to be made with respect to the chemical nature of the factor, and of possible substances *p*-aminobenzoate was found highly active biologically, and to satisfy the chemical requirements. It was later isolated from yeast in the quantities suggested (Blanchard, 1941). Here the interpretation of the mode of action (cf. also Fildes, 1940*b*) preceded chemical characterization of the new nutrient, and its later identification supported the postulated mechanism; it has also been supported by quantitative studies of the course of growth (§ V 1*b*). Data and theories concerning the antibacterial actions of the sulphonamides are more fully discussed in Henry's (1943) review.

(γ) *Acridine derivatives and compounds of related action.* Naturally occurring antagonists to the antibacterial action of acriflavine were found (McIlwain, 1941*b*) to include more than one type of compound. Certain extracts owed their activity to nucleates, which appeared to antagonize the effects of acriflavine by combining with it. Others owed their activity to their content of amino-acids, which were suggested to be the normal products of enzymes inhibited by acriflavine. Valko & Dubois (1944) found the compound to be antagonized also by sodium dodecyl sulphate, and emphasized that a non-toxic ion could antagonize an inhibitory one by phenomena comparable to the ionic exchange known to occur in simpler systems. Such may also be the basis of the antagonism of the action of atebrin on *Bact. coli*, by spermidine and other polyamines (Silverman & Evans, 1943). Similar polyamines antagonize the action of propamidine upon *Lactobacillus casei* and *Streptococcus lactis* (Snell, 1944); other evidence of similarities in the action of propamidine and acriflavine is given by the finding of cross-resistance to their action upon staphylococci (McIntosh & Selbie, 1943). To affect growth, the ionic exchange presumably occurs at components which are important in growth, and it affects, directly or indirectly, metabolic processes of the bacteria. Such interpretations thus have much in common with metabolic hypotheses mentioned above, in the case of acriflavine, and which suggest inhibitor and antagonist to interact at enzymes.

(δ) *Iodinin* (the di-*N*-oxide of a dihydroxyphenazine), a natural antibacterial agent, has also been studied in this respect (McIlwain, 1941*c*, 1943*b*).

These examples have shown the antagonism of drug action by naturally occurring substances to be of value in indicating which of the many natural materials or types of material can be expected to be associated with the action of such inhibitors as are susceptible to antagonism. It thus substitutes a defined procedure for the empiricism of testing all available substances or enzyme systems. The interpretations which it offers of the processes involved are, however, limited and are based upon individual hypotheses which must be further examined by other methods. Instances of such further examination follow.

(a_2) *Observations related to those of drug antagonism*

(α) *Bacterial inhibition by metabolite analogues.* The probability of such inhibition followed from the findings and hypotheses of § a_1 (Fildes, 1940*b*; McIlwain, 1940*a*, 1942*a*). Many substances of importance in bacterial growth were known from straightforward studies of bacterial nutrition, which had been shown to involve a variety of substances similar to those of animal nutrition (cf. Knight, 1936). Analogues $R.SO_3H$ or $R.SO_2NH_2$ corresponding to growth factors $R.COOH$ were first tested in the case of nicotinic acid (R=—C_5H_4N; in the sulphanilamide/*p*-aminobenzoate model, R=—$C_6H_4.NH_2$). The compounds concerned, pyridine-3-sulphonic acid and pyridine-3-sulphonamide, inhibited certain organisms requiring nicotinic acid in growth (McIlwain, 1940*a*, cf. Erlenmeyer, Bloch & Kiefer, 1942; Matti, Nitti, Morel & Lwoff, 1941; Möller & Birkofer, 1942), and in most cases the inhibitions were antagonized by nicotinic acid or its derivatives, which confirmed the nature of the inhibition. Similarly, various α-aminosulphonic acids (R=—$CH(NH_2)R'$) inhibited the growth of several bacteria and were antagonized by α-aminocarboxylic acids (McIlwain, 1941*a*). Special instances of this relationship are those between pantoyltaurine and pantothenate (R=—$CH_2CH_2NH.CO.CHOH.C(CH_3)_2.CH_2OH$; Snell, 1941; Kuhn, Wieland & Möller, 1941; McIlwain, Barnett & Robinson, 1942; McIlwain, 1942*b*, *d*) and pantoyltauramide and pantothenate (McIlwain, Barnett & Robinson, 1942; McIlwain, 1942*d*; Barnett & Robinson, 1942).

Inhibition by substituted thiophenols, diphenyldisulphides, and some selenium and tellurium analogues ($R.SH$, $R.SS.R$, etc.) corresponding or approximating to *p*-aminobenzoic acid ($R.COOH$) were examined by Green & Bielschowsky (1942) and their bacteriostatic action found to be in part related to *p*-aminobenzoate. The thiol, sulphide, disulphide, sulphoxide and sulphone ($R.SH$, R_2S, $R.SS.R$, R_2SO, R_2SO_2) corresponding to pantothenic acid ($R.COOH$) were found to be related to pantothenate in their action (Barnett, 1944). It has also been found that a homologue (of methionine: Harris & Kohn, 1941*a*) and an olefin (related to tryptophan: Fildes, 1941) inhibit bacterial growth through their simi-

larity to bacterial metabolites. These relationships were again applied in preparation of inhibitory analogues of pantothenate (Barnett & Robinson, 1942; McIlwain, 1942*d*). Evidence of the connexion of their effects with the corresponding metabolite is less clear in the latter cases, and also in that of diazine di-*N*-oxides which were designed on the basis of relationship to quinones (McIlwain, 1943*b*, *c*). Substances $R.CO.R'$ related to acids $R.COOH$ have also been found antibacterial: the *p*-aminobenzoate analogues, *p*-aminobenzamide and aminodiphenylketones by Hirsch (1941–2), Kuhn, Möller, Wendt & Beinert (1942) and Auhagen (1942). Wyss, Rubin & Strandskov (1943) found nuclear substitution product of *p*-aminobenzoate to be inhibitory.

Many other analogues of growth-promoting substances have been examined as inhibitors (cf. Woolley, 1944). Not all such analogues, designed on the basis of simple structural similarity, are inhibitory to growth, though substances not inhibiting growth may still inhibit reactions of the organism. It was observed (McIlwain, 1941*a*) that the production of compounds capable of interfering with a given process evidently demanded consideration of properties of which the structural formulae gave only approximate indications. An instance in which a property of this class has been examined in detail is given in the following section. The further characteristics necessary for chemotherapeutic action are discussed later; the essential theoretical link between such instances of bacterial inhibition and the action of classical chemotherapeutic agents was the regarding of drug receptors as enzyme systems or their components, and with this hypothesis it was possible to offer interpretations of many of the pre-existing problems of chemotherapy. These included the specificity of the agents, chemotherapeutic interference, drug resistance and also certain general features of the origin and action of naturally occurring drugs (McIlwain, 1942*a*, 1944*a*).

(β) *Measurement of sulphonamide inhibition and its relationship to acidic dissociation constants of the drugs.* The impression that a particular sulphonamide is specifically associated in its optimal effect with a particular bacterium grew through the introduction, subsequent to that of sulphanilamide itself, of sulphapyridine as an antipneumococcal agent. It is probably without foundation (Green & Parkin, 1942; Wyss, Grubaugh & Schmelkes, 1942); sulphapyridine is also more effective than sulphanilamide against streptococci, but its greater activity widens its range to include the pneumococci. As the drugs are used in practice in the presence of *p*-aminobenzoate, it is of value to compare their activities in terms of the amounts required of them to inhibit bacterial growth in the presence of defined quantities of that substance. This was carried out by Wyss, Grubaugh & Schmelkes (1942) in an extended comparison of the effects of six sulphonamides upon seven bacterial species. The order of activities of the drugs upon different bacteria was found to be essentially the same. The description of the efficacy of an inhibitor, acting by a competitive mechanism, in terms of the ratio of its limiting inhibitory concentration (C_I) to the coincident concentration of the corresponding metabolite (C_M) was proposed by several workers. Such measures of bacteriostatic potency (C_I/C_M; Rose & Fox, 1942), antibacterial indices (C_I/C_M; McIlwain, 1942*b*, *d*) or bacteriostatic constants (C_M/C_I; Wood, 1942) are of value in describing the semi-quantitative phenomena encountered in such investigations. Typical values of C_I/C_M are (Wyss, Strandskov & Schmelkes, 1942; Wood, 1942—values calculated from whose results are given in parentheses): sulphanilamide, with *Bact. coli*, 2000 (1610); with *Staph. aureus*, 4660; sulphapyridine, 450 (100) with *Bact. coli*; with *Staph. aureus*, 416; sulphathiazole, with *Bact. coli*, 27 (36); with *Staph. aureus*, 53. Closer examination shows that the ratio with a given drug and organism is not constant (McIlwain, 1944*b*) but varies with the absolute concentrations of inhibitor and antagonist; upon a probable basis derived from enzyme kinetics, this behaviour would be expected. Various factors peculiar to the conditions of growth of the organism, to its other activities or to the criteria of bacteriostasis adopted, are necessarily also interposed between any system upon which the drugs act, and the growth phenomena observed. These points explain the differing values obtained with a given organism and drug, for C_I/C_M, or for the limiting concentration of drug needed for bacteriostasis in the absence of added antagonist.

Correlations between such values and other properties of the drug cannot, therefore, be expected to be expressed with a high degree of accuracy. This applies to the important correlation which has been found, between the acid dissociation constants of the drugs, and their antibacterial properties. The constants were examined by Bell & Roblin (1942) upon the basis of competitive action between *p*-aminobenzoate and sulphonamides. It was argued that the resemblance between inhibitor and metabolite, upon which the action of the former appeared to depend, was likely to involve the configuration and charge of groupings in the molecules. The configuration was already known to be important: the *o*- and *m*-isomerides of sulphanilamide are not antibacterial. The acidic and basic dissociation constants were chosen as giving measures of the charges of the sulphonamide and amino groupings of a large series of sulphonamides. Except when swamped by the effects of groups elsewhere in the molecule, the basic dissociation constants were similar, and close also to that of *p*-aminobenzoic acid (cf. also Albert & Goldacre, 1942). The acidic constants varied over a considerable range; when plotted against antibacterial activity at *p*H 7, a definite maximum in activity was found to be associated with a pK_a of 6·5. Bell & Roblin deduced from the electromeric characters of groups in sulphonamide molecules a

closely similar value for the dissociation constant expected to be associated with optimal effect, and tested the argument by successfully predicting the approximate dissociation constant and antibacterial activity of *N'*-chloroacetylsulphanilamide. In addition, variation in activity with *p*H was accounted for on theoretical grounds (cf. also, Schmelkes, Wyss, Marks, Ludwig & Strandskov, 1942). One of the synthetic organic chemist's demands of chemotherapeutic theory has thus been satisfied in the present special case, but only after considerable biological investigation; also, the factors promoting antibacterial activity may or may not promote other actions necessary in chemotherapy. Connexion with acidic dissociation constant was established only when sulphonamides of the optimal properties for antibacterial action had already been prepared; the correlation is nevertheless of great value in emphasizing that the properties to be considered in further developments of compounds acting in the manner of sulphanilamide are pharmacological rather than bacteriological, and the demonstration that connexions of this type can be established is an important stimulus to their essaying at earlier stages in the investigation of other agents.

The above arguments apply only to those actions of sulphonamides which are antagonized by *p*-aminobenzoate. The major effects of sulphonamides are so antagonized, though in certain sulphonamides and in analogous compounds, other types of action are encountered (Dorfman & Koser, 1942; Green & Bielschowsky, 1942).

(*γ*) *Antibacterial action of bases.* One of the earliest hypotheses relating to the action of chemotherapeutic agents, which is embodied in receptor theories and was inherent in Ehrlich's use of dyes as chemotherapeuticals, is that the agents act by combination with the affected cells. Such combination has been suggested in the case of bases to be with acidic components of the bacterial cells, as complex salts or by ion exchange (§ *V a*). Both can be expected to depend upon the basicity of the compound concerned. Stearn & Stearn (1924) elaborated the suggestion that combination in the case of basic triphenylmethane dyes was between cations of the bases and acidic components of the parasite. Their main evidence was the increasing activities of the compounds with increase in their estimated strengths as bases, and with increasing *p*H of the solutions in which they were examined. The amino-acridines have recently been examined more fully from this point of view (Albert, 1942; Rubbo, Albert & Maxwell, 1942). The order of antibacterial potency of seven such compounds was roughly the same with five different bacterial species. An arbitrary method of assessing the general antibacterial effect was chosen and measures given for the activities of the compounds, which increased from 1-amino- to 5-amino-acridine. This order was also that of increasing basic dissociation constant. The results have been extended to other members of the series and tend to show that both antibacterial activity and basicity (and also oil-water distribution coefficients) are dependent upon electromeric properties of the substituted molecules; in particular, that type of substitution favouring resonance markedly increased basicity and antibacterial action.

(*b*) *Influence of inhibitors upon the course of growth of parasites*

Section (V *a*) emphasized both the value and limitations of relatively qualitative studies of bacterial growth, in understanding chemotherapeutic problems. Measurement of the course of bacterial growth makes good certain of the deficiencies of those methods. The procedures and description of results in such studies have been adapted from those of general bacteriology. Enumeration of viable organisms is usually by counting colonies found after growth on suitable solid media (cf., however, Lodge & Hinshelwood, 1943). The total growth has been assessed by counting cell numbers, by photoelectric measurement of turbidity, and by O_2 absorption or CO_2 output of cultures.

The course of growth has usually been found to involve lag, logarithmic and stationary phases (Fig. 1; cf. Buchanan & Fulmer, 1928, for a more elaborate classification) and has been characterized by: (1) the value of the stationary population, n_s, obtained from the inoculum of size n_0; (2) the slope of the growth curve during its logarithmic phase, when, if populations n_1 and n_2 are found at times t_1 and t_2, $n_2 = n_1^{K(t_2-t_1)}$. K is then the velocity constant of the rate of growth, obtainable from the above relationship in the form $K = \frac{2{\cdot}3}{t_2 - t_1} \log \frac{n_2}{n_1}$, and is used by Kohn & Harris (1941) to characterize this phase. Its reciprocal is the time required for the population to double during the logarithmic period and this value (T, in minutes) is thus a mean generation time, or reproduction time, and will be used here for describing the acceleration in growth during the logarithmic phase. As normally measured, it is an overall value including any contemporaneous death-rate of the cells. (3) The third value employed is the length of the initial lag phase, formally considered as extending to a time t_l (Fig. 1) when n would equal n_0 had the growth throughout possessed the characters found during the logarithmic phase.

Substances known to be involved in normal bacterial growth have been found to have characteristic effects upon its different phases (Dagley & Hinshelwood, 1938; Lodge & Hinshelwood, 1939, 1943): lowering the concentration of lactose in *Bact. lactis aerogenes* cultures had its major effect upon n_s; of CO_2, upon T; of Mg and amino-acids, upon lag.

(*α*) *Effects of sulphonamides.* Different effects of the drugs upon bacterial growth have been reported

by different workers, but each approach has given valuable evidence of their mode of action. Kohn & Harris (1941) found the main effects of sulphanilamide, sulphapyridine, and sulphathiazole to be upon T; in two distinct media—one simple, the other nutritionally rich, and which afforded different values for T—lag was unaffected both when the total and viable populations were followed, or when inocula of varying sizes were used. The magnitude of the effect upon T in many cases increased with time between the first and sixth hour of action. A result similar except in the last particular was found by Wyss (1941), who studied the mutual interaction of sulphanilamide and *p*-aminobenzoate upon the same organism. He interpreted his results in terms of the hypothesis that the two compounds compete for an enzyme whose activity depends upon its reaction with *p*-aminobenzoate and conditions the rate of growth of the culture. The rate of growth was found to vary with the concentrations of growth-promoting and growth-inhibiting substances in a manner which (with a certain limitation: Gaddum, 1943) agreed with an expression (Lineweaver & Burk, 1934) derived from enzyme kinetics. The finding did not necessarily demonstrate the drug and antagonist to combine with an enzyme (Wyss, 1941); similar equations could be derived on the basis of adsorption hypotheses, or receptor mechanisms (Gaddum, 1937, 1943), but neither of these, without further assumptions (such as the receptors being enzymes; McIlwain, 1942*a*), could explain the connexion between substances and living processes. There is also other evidence of the action of sulphonamides upon metabolic processes.

Davies & Hinshelwood (1943), investigating the effects of sulphonamides upon *Bact. lactis aerogenes* in simple media, found the action of the drug to vary with treatment of the inoculum, especially its age and previous contact with sulphanilamide. At an intermediary age when lag was minimal, results similar to those of Kohn & Harris (1941) were obtained, in that inhibition increased with time, though a later, rapid, phase not reported by those workers was sometimes found. Davies & Hinshelwood (1943) regard as most typical the effects of the drugs upon young organisms, which in their normal growth show lag and a single logarithmic phase. In the presence of sulphanilamide these organisms show increased lag and two logarithmic phases: a slow one succeeded by a more rapid one (II, Fig. 1). This the authors interpret as due to two modes of growth of the organisms, differing in susceptibility to the inhibitor. The mode of shortest lag is necessarily that of normal growth; it also has the smaller T and is supposed to be the more susceptible to sulphanilamide, which, by increasing its lag to beyond that for the second mode, allows this process to manifest itself. Increasing sulphanilamide concentrations increased the T of both processes. The mechanism was also supported by training experiments. Values of n_s were found to increase with relatively low concentrations of sulphanilamide, but to fall with higher concentrations; this may be associated with *p*H changes. Excess *p*-aminobenzoate annulled all the effects of sulphanilamide and restored the original mode of growth; the effects of intermediate concentrations were not reported. Wyss's experiments, which were concerned with such concentrations, leave open the possibility of the later supervention of a type of growth corresponding to the later more rapid phase of *Bact. lactis aerogenes*. The organisms and other conditions also differed in the two investigations.

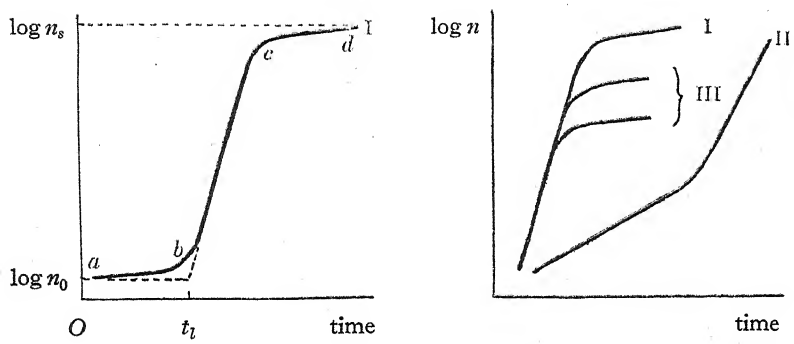


Fig. 1. Types of normal (I) and inhibited (II, III) growth. Ordinate: logarithm of a measure of cell numbers or population density (n_0, initial and n_s, stationary populations). Curve I: *ab*, lag phase (of length Ot_l); *bc*, logarithmic phase; *cd*, stationary phase.

Increase of T (to T_i while K alters to K_i) in the presence of sulphonamides was expressed by Davies & Hinshelwood (1943) as a ratio, T/T_i, which decreased with increasing concentrations of the inhibitor but was not found to fall to zero. The effect of the composition of the medium of growth, upon the type of curve obtained by plotting the related ratio K_i/K against the logarithm of sulphanilamide concentrations, was studied in detail by Kohn & Harris (1941). They considered that the basic type of action in a simple medium was of

linear relationship between the factors (I, Fig. 2); certain of the results of Davies & Hinshelwood (1943) also show this relationship, and types of relationship obtained with other substances are discussed by Poole & Hinshelwood (1940). The effects with sulphonamides in rich media, for example in one containing peptone, were more complex (III, Fig. 2; Kohn & Harris, 1941), and showed that the medium antagonized the drug. Attempts to obtain this effect with constituents of the peptone showed that methionine reproduced it to some extent (II, Fig. 2; Harris & Kohn, 1941*a*; Kohn & Harris, 1943; cf. also Bliss & Long, 1941). A shift similar to that from curve I to curve II was given also by *p*-aminobenzoate, but whereas the effect of methionine was limited (i.e. capable of antagonizing limited concentrations only of sulphanilamide), *p*-amino-

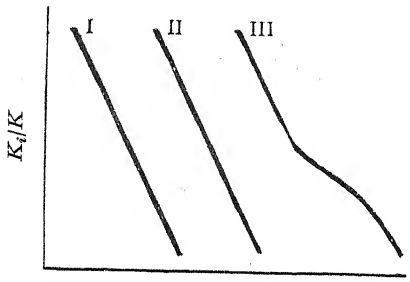


Fig. 2. Effect of composition of medium upon growth in the presence of sulphanilamide. Ordinate: ratio of growth constants of inhibited to normal growth.

benzoate was effective against a much larger range of sulphanilamide concentrations. The material in peptone associated with curve III could not, however, be methionine or *p*-aminobenzoate alone, as its effect was qualitatively different, the curve no longet being linear, and its form varying with differenr sulphonamides in a way in which those associated with other antagonists did not. These effects were to some extent similar to those of acriflavine antagonism (§ Va_1), and Kohn & Harris's interpretation of them was also in terms of consecutive metabolic processes and involved the hypothesis that substances which antagonized the drugs were related to substances normally made by the organisms, but whose synthesis was disturbed by the drug. Growth of the organism was considered to depend upon materials for whose synthesis *p*-aminobenzoate was necessary, but to different degrees: sulphanilamide inhibition first, in relatively low concentrations, leading to lack of methionine, but in higher concentrations to lack of the substances in peptone, and in still higher concentrations to lack of other substances, unknown as nutrients, but capable of synthesis in the presence of *p*-aminobenzoate.

Certain purines also antagonized sulphonamide inhibition (Harris & Kohn, 1941*b*; Kohn & Harris, 1943) but did not account for the full effect of the peptone. Additional antagonistic effects have been found to the actions of sulphapyridine and sulphathiazole upon dysentery bacilli. Here nicotinic acid (which to those organisms is a necessary nutrient) also contributed to antagonism, but the effect was independent of, and was not substituted by, *p*-aminobenzoate; it presumably concerns a different metabolic series (Dorfman & Koser, 1942). Synergism between sulphanilamide and other compounds, especially urethane, has been studied in detail by Johnson, Eyring & Kearns (1943).

(β) *Pantoyltaurine.* This substance (cf. § Va_2) was prepared with the intention of its interfering with pantothenate metabolism. The main antagonist to its antistreptococcal action was found to be pantothenate; other substances of serum may have small effects but were not sufficiently well characterized to be studied by the methods of Kohn & Harris. An alternative method to the study of its effects was afforded by the inhibited organisms requiring pantothenate in their normal growth (McIlwain, 1944*b*). Growth of β-haemolytic streptococci was measured in a complex medium whose pantothenate content could be adjusted by direct addition of the compound. Growth with excess pantothenate showed a short lag and a well-marked logarithmic phase; increasing quantities of pantoyltaurine in the presence of pantothenate were found to have their main effect upon the logarithmic period, which divided into at least two phases: an initial one, much inhibited by pantoyltaurine, and a later, more rapid one, with T only a little greater than that of uninhibited cultures; n_s was little affected. The result was similar to curve II of Fig. 1. The initial presence of more pantothenate prevented all the effects of pantoyltaurine, which thus appeared to act by limiting the use of the growth factor by the organisms; the structures of the two substances, and other evidence, favoured competitive inhibition as mechanism of the interaction. When, however, the use of pantothenate by the organisms was limited by decreasing the quantity added to the basal medium, in absence of the inhibitor, a different result was obtained. Lag and T were increased only slightly; the main effect (as in III, Fig. 1) was upon n_s, which below 5×10^{-9} M pantothenate was proportional to the quantity of pantothenate added.

The interpretation of these results was based upon the knowledge that, although lowered concentrations of pantothenate did not have their major effect upon T, limitation of other nutrients could have this effect. It was therefore considered that the effect of pantothenate upon growth was indirect, through its participating in the production of secondary substances, and that pantoyltaurine controlled growth through inhibiting their formation. This is in agreement with metabolic phenomena discussed below. The rate of growth during the initial phase

possessed characters similar to those found in sulphanilamide-limited growth of *Bact. coli* by Wyss (cf. the foregoing section) and therefore supported the suggestion. Transition to the more rapid phase of growth is open to interpretations similar to those of Davies & Hinshelwood (1943): the existence of two modes of growth differing in pantoyltaurine susceptibility; the formation of substances necessary to growth; and also the reaching of a critical population level.

Little data is available concerning the detailed effects of other chemotherapeuticals upon phases of bacterial growth. Inhibitors other than chemotherapeutic agents have been studied in this respect by Dagley & Hinshelwood (1938), Poole & Hinshelwood (1940) and Spray & Lodge (1943). Penicillin preparations may affect *T* (Foster & Wilker, 1943) or lag (Fisher, 1943).

(*c*) *Reference of growth effects to events in the parasite*

(α) *Sulphonamides.* Sulphanilamide is absorbed by bacterial cells, though the amount remaining after washing is not greater than are those of therapeutically inactive isomerides (Feinstone, Williams & Florestano, 1942). It does not displace from resting cells an appreciable proportion of the *p*-aminobenzoate which they contain (McIlwain, 1944*c*). The equilibria and competitive interaction found between the two compounds in bacterial growth thus extend to only a fraction of the *p*-aminobenzoate, or to a particular phase in its metabolism, e.g. to its combination with cell constituents.

The belief that the sulphonamides might act by metabolic interference preceded both the identification of their antagonists and the detailed examination of their effects upon growth, and led to various suggestions concerning their mode of action, which have been discarded as inconsistent with later information. Among investigations founded upon knowledge of sulphonamide antagonists is that of the inhibition by the drugs of the oxidation of *p*-aminobenzoate by peroxidase (Lipmann, 1941). Oxidation of phenols by tyrosinase is also inhibited (Baur & Rüf, 1942). No effect of sulphonamides upon certain dehydrogenations in organisms susceptible to the drugs was found by Mellon & Bambas (1937), and little effect upon general respiration by Ely (1939) and Hirsch (1942). Effects reported by Sevag & Shelburne (1942) are given under certain conditions by therapeutically inactive isomerides (Wyss, Strandskov & Schmelkes, 1942), but are probably related to the action of the drugs (Henry, 1943). A more specific effect is given by the observation of Fox (1942) that, during growth of *Bact. coli* in the presence of sulphanilamide, a new diazotizable substance, not *p*-aminobenzoate, accumulated in amounts which precluded its being produced from the drug alone.

(β) *Pantoyltaurine.* This compound, like sulphanilamide, acted upon growing cultures only after a time lag and displaced little if any pantothenate from preformed cells of streptococci whose growth it inhibited (McIlwain, 1944*c*). These organisms required the addition of preformed pantothenate for their growth. When growing cultures were examined, much more pantothenate was found to disappear from the culture fluid than could be explained in terms of assimilation to the cells. This was found to be due to a metabolic process, occurring also in resting cells during glycolysis. Pantoyltaurine strongly inhibited the former process, though it had little effect upon glycolysis. The inhibition of pantothenate metabolism was brought about by concentrations of pantoyltaurine about one-tenth of those affecting growth. The reaction was thus sufficiently sensitive to the inhibitor to explain its effects upon growth; it concerned pantothenate, whose connexion with the process of inhibition is established, and glycolysis which is an important reaction to the organisms inhibited. The abilities of various compounds related to pantoyltaurine to inhibit growth of streptococci were correlated with their abilities to inhibit the metabolic process. The process was then considered to be the basis of the bacteriostatic and chemotherapeutic effects of the drug. To explain inhibition of growth of the streptococci, by a process which preserved the cognate growth essential, pantothenate, it was suggested that the functioning of pantothenate involved its undergoing a cycle of changes, which was interrupted by pantoyltaurine, and of which at least one phase was relatively unstable in the reaction mixture examined.

(γ) *Other agents.* Examination of the diamidines as chemotherapeutic agents was based upon an effect upon the metabolism of the host (hypoglycaemia) which is not the mode of action of the compounds in chemotherapy (Yorke, 1940), though they may affect glucose metabolism in the parasite. They inhibit the Pasteur effect in various cells (Dickens, 1939). Of substances which are not chemotherapeutic agents, several inhibitors which are antibacterial have defined metabolic effects though not all have been fully examined in bacteria themselves. Iodoacetate, fluorides (Wiggert & Werkman, 1939), and narcotics (Johnson, Brown & Marsland, 1942) are among those which have been so examined (cf. also Clark, 1937 and Bernheim, 1942). The azochloramide potentiation of sulphonamides (Neter, 1942 *a*, *b*) has been referred to a chemical mechanism (Schmelkes & Wyss, 1942). Penicillin-B or notatin is itself an enzyme and owes its antibacterial properties to the hydrogen peroxide produced in its oxidation of glucose (Coulthard, Michaelis, Short, Sykes, Skrimshire, Standfast, Birkinshaw & Raistrick, 1942; Roberts, Cain, Muir, Reithel, Gaby, van Bruggen, Homan, Katzman, Jones & Doisy, 1943; Hirsch, 1943). An enzyme capable of chemotherapeutic action is discussed in § V 3.

2. *Effects upon viability*

Normal bacterial cells show definite rates of death (Wilson, 1922), so that conditions which delay their growth, or sufficiently lower its rate, can themselves result in the death of cultures (cf. Lodge & Hinshelwood, 1939). In the host, other factors increase the death-rate of the parasite, but there is less evidence that the chemotherapeutic agents at present successful in bacterial infections directly do so. The present discussion is therefore brief, though the application of bactericidal agents as disinfectants is not sharply differentiated from chemotherapy. Bactericidal effects of the sulphonamides have been reported only under certain limited conditions which are different from those of chemotherapy. Their normal action, in stopping growth but not causing the irreversible changes associated with bactericide, has been related to the finding that sulphanilamide, though competing with *p*-aminobenzoate in processes concerned with the growth of cultures, does not displace it from preformed cells (§ V 1 *c*).

In an instance in which the characteristics of both bactericidal and bacteriostatic actions of the same agent (crystal violet) have been examined, the two actions have been found to be distinct in mechanism (Hoffmann & Rahn, 1944). Of factors causing the irreversible changes involved in bactericide, Albert (1942) has emphasized surface activity in the agent. Baker, Harrison & Miller (1941 *a*) state that bactericidal action and surface-tension depression of compounds are not closely correlated; other types of agent are bactericidal. Many of these can, however, simply be regarded as causing gross chemical or physical changes in the organism, and the action of the group of surface-active compounds shows certain characteristic features. Investigation of the type of material antagonistic to them (Baker *et al.* 1941 *a*, *b*) has afforded results which can be interpreted as ion exchange between the inhibitor and antagonist (Valko & Dubois, 1944). This is related to their functioning through the suggestion that the attachment of the inhibitory ions to enzymes inactivates them, while attachment of the antagonistic ion does not.

The natural antibacterial agents, gramicidin and tyrocidin, have many characters in common with detergents (Dubos & Hotchkiss, 1942), though gramicidin is in some cases primarily bacteriostatic and has a more specific action. The agents are important in exemplifying a biological rationale in the preparation of substances of biological activity (Dubos, 1939*a*; and below). The efficacy of preparations containing them, against pneumococcal infections in mice, has been demonstrated (Dubos, 1939*b*), but occurs only if they are injected in the same site as the cocci (Robinson & Graessle, 1942). As the action of the bactericidal component is prevented by serum while that of the bacteriostatic component is not, it is probable that here also the action of the agent upon the cocci *in vivo* is a bacteriostatic one. Effects upon metabolic processes in the organisms affected have been reported (Dubos & Hotchkiss, 1942). These and other agents exhibit a species-specificity in their action upon micro-organisms, which is correlated with Gram-staining (Dubos, 1939*a*, 1942) and hence also with certain constituents of the cell whose chemical nature has to some extent been specified (Henry & Stacey, 1943).

3. *Effects upon products of the parasite*

Therapeutic effects through the administration of substances which affect bacterial products have long been studied in immunology. That subject has a formal distinction from chemotherapy in that its reagents may be produced by the host itself as a normal reaction to parasitism; the basis of the separation, however, is questionable (cf. Dubos, 1941). Antibodies used in therapy are commonly prepared in organisms other than the host to which they are applied; in this respect they are similar to quinine. They are more complex than orthodox chemotherapeuticals, but so also are the anthelmintic proteases. Moreover, antibodies have been prepared *in vitro* from separated plasma proteins (Pauling & Cambell, 1942). The production of quinine and of other natural antibiotic agents may also be conditioned by parasitism (cf. McIlwain, 1944*a*). Immunological reagents thus afford examples of chemotherapeuticals, some of which act by combination with bacterial products which are offensive to the host.

Of orthodox chemotherapeutic agents, sulphonamides have been believed to act through destruction of bacterial toxins, but the possible scope of such action has been limited by later investigation (Zahl, Hunter & Cooper, 1944). Sulphonamide action has also been considered to be due to an effect upon a different product related to pathogenicity, namely, the capsules which, especially in pneumococci, are concerned in their protection from bactericidal agencies of the host. Clear demonstration of the bacteriostatic action of the sulphonamides *in vitro* in the absence of leucocytes afforded evidence against such an action being the primary effect of the drugs (McIntosh & Whitby, 1939). A chemotherapeutic agent which acts in this way has, however, been obtained by the novel procedure of Dubos (1939–40; Avery & Dubos, 1931), who searched for natural agents capable of destroying the capsular material of pneumococci, by presenting it as substrate to soil and other likely sources of organisms capable of the required reaction. Bacteria which could hydrolyse the polysaccharides of pneumococcal capsules were isolated, and from them an enzyme separated which itself was capable of the hydrolysis. The enzyme preparation protected mice from pneumococcal infec-

tion (Avery & Dubos, 1931) and thus had the additional virtue, not susceptible to intentional introduction by the method of preparation, of some suitable pharmacological characters. It was extremely specific, its action being confined to the type of pneumococcus whose capsular material (which is responsible for reactions of pneumococcal typing) was used in preparation of the reagent. It is probable that not only was the organism producing the enzyme isolated by selectively feeding the polysaccharide, but that also its metabolism became adapted to the unusual substrate: for continued production of the enzyme, cultivation in the presence of the polysaccharide was necessary.

VI. DRUG AND HOST

The main types of interaction to be noted between these components of the chemotherapeutic system are: (1) distribution of the drug in the host, its absorption and excretion; (2) chemical changes of the drug *in vivo*, which may result in its activation or inactivation; (3) effects of the drug upon the host, including toxic action or possible stimulation of the host's defences. The subject thus consists largely of the pharmacology of chemotherapeutic agents, and its adequate study is to be made in terms of pharmacology and physiology.

A relatively simple method of observing the overall effects of the above factors is to determine *in vitro* the effects upon the parasite concerned, of the blood or relevant tissues of the host, at varying periods after administration of the drug. The length of time during which an antibacterial effect can be maintained is thus determined, but the extent to which this is governed by absorption, excretion, activation, or antagonism remain uncertain. Purely microbiological methods can in some cases be used to estimate the concentrations of both drug and antagonists in tissues and fluids. The varying effects can then be assessed with greater probability; the method remains an indirect one with certain valuable applications (McIlwain & Hawking, 1943) but unsuitable for allocating to definite factors the differences between related compounds. All interactions between drug and host can vary to a considerable extent in different hosts (cf. Marshall, 1939, 1941 *a*, *b*; McIlwain, 1943 *e*). Though knowledge of certain general trends in such factors is part of the stock-in-trade of a pharmacologist or chemotherapist, the examination of several animal species prior to clinical application is necessary. The manner of administration of the drug is a factor under experimental control which conditions its overall effect. Local application to wounds involves special problems of absorption and toxicity.

(1) Considering the first of the factors enumerated above, much data is available concerning the absorption, distribution and excretion of sulphonamides. Shannon's (1943) account is valuable in giving quantitative values for distribution in several sites, and for the glomerular filtration rate and rate of reabsorption by the renal tubules, of many sulphonamides and related compounds. Only when these and other processes are assessed separately can any extensive correlation between the chemical and pharmacological properties of substances be expected. The processes were found to be greatly conditioned by the degrees of ionization of the agents at the *p*H of the tissues; sulphanilamide itself and several derivatives of high pK_a were widely distributed while those of low pK_a tended to have the limited distribution of Na or Cl ions. Excretion and reabsorption was also conditioned by these properties, but by others in addition. The *o*- and *m*-isomerides of sulphanilamide were not very different from it in this respect. Access to the brain was specially conditioned. Drugs may be 'bound' in the body in non-diffusible and chemotherapeutically less active forms (Davis, 1943).

(2) The importance of chemical change of the drug in the host, in producing active agents from much less active compounds administered, was demonstrated in studies of the arsenicals and encountered early in bacterial chemotherapy: prontosil and certain other sulphonamides were found to be active only through their breakdown *in vivo* to sulphanilamide itself. It is possible to take advantage of such interaction between host and drug to secure a gradual supply of an active agent without frequent dosage, or a distribution of the agent different from that obtainable by its direct administration. The reactions concerned may be reproducible *in vitro*, and study of their nature and velocities are necessary to give a secure basis to their therapeutic use.

Other actions of host upon drug may be less desirable chemotherapeutically. Many drugs are converted to less active forms by the host (cf. Bernheim, 1942): some sulphonamides to less soluble acetyl derivatives which have the additional disadvantage of disturbing the host mechanically, through their crystallization *in vivo*. Such separation, and its locality, is susceptible to a certain degree of analysis in terms of the physical and chemical properties of the products, especially of their solubility at various *p*H values, but involves also such physiological characters as their rate of renal clearance. Acetylation of the sulphonamides is among the processes which have been reproduced in tissue preparations *in vitro*; the factors concerned include general biochemical ones such as the supply of pyruvate or acetate.

(3) The type of pharmacological investigation of chemotherapeutic agents most immediately relevant to their use is the assessment of their toxicity, preferably in a defined form such as the LD_{50} and connected with the use of the drug in an expression such as the chemotherapeutic index. The results vary markedly with the manner of administration of the drug. The toxicology of a single type of agent

can be a considerable study: among the sulphonamides, the major toxic effect can be through mechanical blocking by the acetyl derivative, inhibition of intestinal symbionts, more direct effects upon blood constituents, drug fever, and many other circumstances (Hawking, 1937; Marshall, 1939, 1941 *a*, *b*; van Dyke, 1943). Little connexion between toxicity and other properties of the drug, useful in planning new agents, can thus be expected without specific studies which refer the actions of the substances to particular sites and processes in the host.

VII. PARASITE AND HOST

The necessity for study of parasitology as an aspect of chemotherapy has become greater with the finding that many agents act chemotherapeutically only in conjunction with the defences of the host; the following points emphasize this.

Most experiments performed with the intention of developing chemotherapeutic agents are carried out in different hosts, and in some cases with different parasites, from those constituting the system against which chemotherapy is desired. The use of such indirect experimental arrangements makes it necessary to understand the differences, in different animal species, of factors conditioning drug action. Straightforward studies of the course of infection have shown the host's antibacterial agencies to include non-specific inhibitory or toxic substances; immunological interaction, and the reticulo-endothelial system. Apart from variations in these factors, comparative studies in chemotherapy include the location of the parasite and drug in different hosts; the reactions of the hosts to the drug in absorption, excretion and chemical change; the hosts' contents of substances with specific relationships to the parasite and drug.

The relationship between host and parasite can be an extremely specific one, though there is much variation in this respect. In some cases experimental infection, but without the characters of the human disease, may be secured. The relation between host and parasite is, however, open to some extent to experimental modification: by passage, a parasite can be increased in virulence with respect to a particular host and induced to infect organisms not normally susceptible to it, though there are limitations to such processes. When the limitations are such that a suitable infection cannot be produced in an experimental animal by a parasite whose study is desired, chemotherapeutic tests have been carried out upon related infections, as in the canary testing of antimalarials. The host-parasite relationship is also open to modification by experimental change in the host; selection has yielded groups of mice susceptible and resistant to mouse typhoid, whose properties were conditioned genetically and correlated with leucocyte numbers (Gowen & Calhoun, 1943). The technique can thus afford hints of the type of interaction between host and parasite as well as providing a range of testing conditions. The location of a parasite in the host often conditions chemotherapy; the factors conditioning elective localization of infections are not fully understood, but are considered to include the tissue content of growth-promoting and growth-inhibiting agents as well as anatomical factors such as the course of the lymphatics.

VIII. DRUG, PARASITE AND HOST

The aspects considered below are mainly those which afford evidence for analysis of the chemotherapeutic system and reference of its phenomena to events in component systems. The most important of the simpler systems has been in all cases that of drug and parasite. Their *in vitro* interaction has never sufficed to explain phenomena *in vivo* with any completeness, but has afforded the most important standard of reference in elucidating the additional types of action occurring in the complete system.

Polysaccharidases. The polysaccharidases of Dubos afford an instance in which the additional type of action has been clearly established. *In vitro*, they are neither bactericidal nor bacteriostatic, but the characteristic disappearance of capsular material from susceptible pneumococci which they cause *in vitro* is also reproduced *in vivo* during successful chemotherapy. Strains resistant to the agents *in vitro* are also refractive in mice. The disappearance of capsular material *in vivo* is correlated with the majority of the bacteria becoming susceptible to phagocytosis, while untreated or resistant cocci do not undergo phagocytosis (Dubos, 1939–40). Evidence is thus afforded for the particular interaction of drug and parasite which is preliminary to the chemotherapeutic response, and for the component of the host which participates in the overall effect.

Bacteriostatic agents. The sulphonamides may be considered in detail. They have no effects upon leucocytic or phagocytic activities; upon the speed of production, or the quantity or quality of production, of immune bodies (McIntosh & Whitby, 1939). Their concentrations *in vivo* during successful chemotherapy are under relevant *in vitro* conditions bacteriostatic only. Action of the drugs under both circumstances is prevented by *p*-aminobenzoate and certain other substances (Selbie, 1940; Martin & Fisher, 1942) with differences which can be understood in terms of the metabolism of the antagonists by the hosts (McIlwain, 1942*c*). A specific synergistic effect with sulphonamides upon bacteria, observed *in vitro*, is reproduced in therapy (Neter, 1942 *a*, *b*; Schmelkes & Wyss, 1942) and is also understandable in terms of *p*-aminobenzoate. The varying effectiveness of the sulphonamides in different types of lesion has in some cases been explained in terms of antagonism of their bacterio-

static action, which can be reproduced outside the host (Lockwood, 1941; Boroff, Cooper & Bullowa, 1942). Their inability, in clinical use, to check initial multiplication of susceptible bacteria and some local infection, is correlated with their delayed action upon the growth of cultures (McIntosh & Whitby, 1939) and again with their inability to displace *p*-aminobenzoate from preformed bacteria. Resistance to the drugs on the part of organisms normally susceptible to them occurs both *in vivo* and *in vitro* in respect to the bacteriostatic effect, to degrees often correlated (e.g. Petro, 1943). Resistance is in some cases further correlated with the ability of the strains to synthesize specific antagonists to the drugs, such as *p*-aminobenzoate (Landy, Larkum, Oswald & Streightoff, 1943). There is thus cumulative evidence of many types for the chemotherapeutic response being due to a bacteriostatic effect upon the parasite, and this being connected with particular substances or processes in the susceptible organisms.

Pantoyltaurine is also bacteriostatic, and not bactericidal, *in vitro*, and bacteriostatic concentrations are attained during chemotherapy. Antagonism of the bacteriostasis is effected *in vitro*, with considerable specificity, by pantothenate and reproduced *in vivo* by comparable concentrations (McIlwain & Hawking, 1943). The resistance of certain organisms is connected with their pantothenate synthesis (McIlwain, 1943*d*). An interaction between pantothenate and pantoyltaurine occurs not only *in vivo* and in the system of drug and parasite but also in a specific metabolic reaction of the parasite, and characters of the latter reaction can explain certain peculiarities of the interaction in growing cultures (cf. § V 1 *b*). The main *in vitro* effect of penicillin, in concentrations effective *in vivo*, is also bacteriostatic (Abraham, Chain, Fletcher, Gardner, Heatley, Jennings & Florey, 1941).

Bacteriostasis alone can clearly afford a basis for action related to chemotherapy. In 12 hr. an infecting inoculum of many pathogens could *in vitro* increase 10^{12}-fold as a result of simple fission at about each 20 min. Bacteriostasis is probably not complete and operates after a short time lag, but reduction of the rate of growth so that division occurs at about each 40 min. after a lag of 3 hr. would still mean that after 15 hr. an agent causing such a change would have limited to one-millionth the number of parasites which would have been present in its absence. The bacteriostatic chemotherapeuticals, however, do not merely delay the course of infection but by a relatively short treatment can permanently prevent its development from a given inoculum. Such a drug can protect the host from numbers of organisms greatly in excess of the minimum necessary to cause death in untreated animals. This result might not be expected on the basis of bacteriostatic action, but the following points have also to be considered. First, the parasites' inherent death-rate can be expected to be increased in the host. Though in pathogenesis the rate remains less than their normal rate of growth it may be greater than their growth rate in the presence of the agent. Secondly, the development of the host's defences requires the presence of foreign organisms; with pathogens, such presence is associated with their growth and production of agents deleterious to the host. In delaying this the bacteriostatic compound can provide greater reaction by the host to a smaller number of bacteria; immunity frequently follows chemotherapy with the sulphonamides. Whether the type of interaction between chemotherapeutically successful bacteriostatic agents, and susceptible parasites, must fulfil immunological criteria or have more specific relationship to phagocytosis than is inherent in growth inhibition, is not established. Such has been suggested, for example, by Stacey & Schlüchterer (1939). The majority of bacteriostatic agents cause morphological distortion before growth is finally inhibited, but without markedly specific characters and apparently independently of their chemotherapeutic value.

There thus remain many gaps in the data connecting chemotherapeutic phenomena with events in simpler systems, but their analysis can be carried sufficiently far to show chemotherapy to be developing from an empirical subject to one with characteristic problems, laws, and types of behaviour.

IX. SUMMARY

Higher organisms and micro-organisms exhibit many types of mutual action and association, and that type of interaction in which the micro-organism becomes a parasite is normally prevented by the potential host. If parasitism is established its course may be impeded by administration of substances to the host; interaction in the resulting system of drug, parasite and host then constitutes chemotherapy.

The interactions of successful bacterial chemotherapeuticals with the parasites, apart from the host, show characters which in most cases afford a basis for their effects *in vivo*. Their action may be to retard bacterial growth, as is the case with sulphonamides. They may stop the use by the parasite of substances which protect it from the host, as do the antipneumococcal polysaccharidases; or they may react with bacterial products which are injurious to the host, as do antitoxins. Such actions are susceptible to further analysis. The effects upon growth have been found to be associated with particular substances whose nature in several cases has suggested the mode of action of the inhibitor. Hypotheses based upon such findings have enabled certain new types of agents to be prepared, and reasons given for their structural specificity. Processes of the parasite, which are affected by chemotherapeuticals, have also been characterized by following the influence of the agents upon the course of growth of the parasite. A beginning has been made to specifying such influence in terms of effects upon isolated metabolic processes and other events in the parasite.

Interaction of the host with both parasite and antibacterial agent or potential agent condition the issue in

chemotherapy. The concentrations of effective agent reached at different times in different parts of the host depend upon characters which are open to experimental modification, such as the mode of administration of the agent, but also upon many which are not. The latter may be accommodated by suitable choice or design of agent and include its physiological characters such as absorption, excretion and reabsorption, and biochemical ones such as its activation or inactivation by substances or processes of the host. It is necessary to study these factors, and other undesirable reactions with the host, before correlation of chemotherapeutic efficacy with structure of the agent can be expected. Events in the complete chemotherapeutic system have been referred to properties of its components by observations which include: correlation of the actions of an agent upon the parasite *in vivo* and *in vitro*; the reproduction *in vitro* of synergism and antagonism observed in chemotherapy; the finding of *in vitro* phenomena associated with drug resistance *in vivo*; and the type of change in the host associated with success of chemotherapy or with normal recovery from, and resistance to, infection.

X. REFERENCES

ABRAHAM, E. P., CHAIN, E., FLETCHER, C. M., GARDNER, A. D., HEATLEY, N. G., JENNINGS, M. A. & FLOREY, H. W. (1941): *Lancet*, **2**, 177.

ALBERT, A. (1942): *Lancet*, **2**, 633.

ALBERT, A. & GOLDACRE, R. (1942): *Nature, Lond.*, **149**, 245.

ANDREWES, C. H., KING, H. & ENDE, M. VAN DEN (1943): *J. Path. Bact.* **55**, 173.

AUHAGEN, E. (1942): *Hoppe-Seyl. Z.* **274**, 48.

AVERY, O. T. & DUBOS, R. J. (1931): *J. Exp. Med.* **54**, 73.

BAKER, Z., HARRISON, R. W. & MILLER, B. F. (1941*a*): *J. exp. Med.* **73**, 249. — (1941*b*): **74**, 611, 621.

BARNETT, J. (1944): *J. Chem. Soc.* p. 5.

BARNETT, J. W. & ROBINSON, F. A. (1942): *Biochem. J.* **36**, 364.

BAUR, E. & RÜF, H. (1942): *Helv. chim. Acta*, **25**, 523.

BELL, P. H. & ROBLIN, R. O. jr. (1942): *J. Amer. Chem. Soc.* **64**, 2905.

BERNHEIM, F. (1942): *Interaction of Drugs and Cell Catalysts*. Minneapolis: Burgess Publ. Co.

BLANCHARD, K. C. (1941): *J. Biol. Chem.* **140**, 919.

BLISS, E. A. & LONG, P. H. (1941): *Bull. Johns Hopkins Hosp.* **69**, 14.

BOROFF, D. A., COOPER, A. & BULLOWA, J. G. M. (1942): *J. Immunol.* **43**, 341.

BUCHANAN, R. E. & FULMER, E. I. (1928): *Physiology and Biochemistry of Bacteria*, **1**. Baltimore: Williams and Wilkins.

CANNON, P. R. (1940): *Ann. Rev. Physiol.* **2**, 387.

CLARK, A. J. (1937): *Heffter's Handbuch der experimentellen Pharmakologie*, Suppl. vol. **4**. Berlin: Springer.

COULTHARD, C. A., MICHAELIS, R., SHORT, W. F., SYKES, G., SKRIMSHIRE, G. E. H., STANDFAST, A. F. B., BIRKINSHAW, J. H. & RAISTRICK, H. (1942): *Nature, Lond.*, **150**, 634.

DAGLEY, S. & HINSHELWOOD, C. N. (1938): *J. Chem. Soc.* p. 1930.

DAVIES, D. S. & HINSHELWOOD, C. N. (1943): *Trans. Faraday Soc.* **39**, 431.

DAVIS, B. D. (1943): *J. Clin. Invest.* **22**, 753.

DEWING, T., GRAY, W. H., PLATTS, B. C. & STEPHENSON, D. (1942): *J. Chem. Soc.* p. 239.

DICKENS, F. (1939): *Biochem. J.* **33**, 2017.

DORFMAN, A. & KOSER, S. A. (1942): *J. Infect. Dis.* **71**, 241.

DUBOS, R. J. (1939*a*): *J. Exp. Med.* **70**, 1. — (1939*b*): **70**, 11. — (1939–40): *Harvey Lect.* **35**, 223. — (1941): *Chemotherapy*. Pennsylvania: Univ. Press. — (1942): *Ann. Rev. Biochem.* **11**, 659.

DUBOS, R. J. & HOTCHKISS, R. D. (1942): *Trans. Stud. Coll. Physicians, Philad.* 4th ser. **10**, 11.

VAN DYKE, H. B. (1943): *Ann. N.Y. Acad. Sci.* **44**, 477.

EHRLICH, P. (1909): *Ber. dtsch. chem. Ges.* **42**, 17.

ELY, J. O. (1939): *J. Bact.* **38**, 391.

ERLENMEYER, H., BLOCH, H. & KIEFER, H. (1942): *Helv. chim. Acta*, **25**, 1066.

FEINSTONE, H., WILLIAMS, R. D. & FLORESTANO, H. J. (1942): *Proc. Soc. Exp. Biol., N.Y.*, **49**, 234.

FERGUSON, J. (1939): *Proc. Roy. Soc.* B, **127**, 387.

FILDES, P. (1940*a*): *Brit. J. Exp. Path.* **21**, 67. — (1940*b*): *Lancet*, **1**, 955. — (1941): *Brit. J. Exp. Path.* **22**, 293.

FISHER, A. M. (1943): *Bull. Johns Hopk. Hosp.* **73**, 343.

FOSTER, J. W. & WILKER, B. L. (1943): *J. Bact.* **46**, 377.

FOX, C. L. jr. (1942): *Proc. Soc. Exp. Biol., N.Y.*, **51**, 102.

GADDUM, J. H. (1937): *J. Physiol.* **89**, 7*P*. — (1943): *Trans. Faraday Soc.* **39**, 323.

GOWEN, J. W. & CALHOUN, M. L. (1943): *Proc. Nat. Acad. Sci., Wash.*, **29**, 114; *J. Infect. Dis.* **73**, 40.

GREEN, H. N. (1940): *Brit. J. Exp. Path.* **21**, 38.

GREEN, H. N. & BIELSCHOWSKY, F. (1942): *Brit. J. Exp. Path.* **23**, 1, 13.

GREEN, H. N. & PARKIN, T. (1942): *Lancet*, **2**, 205.

HAGER, G. P. & GRUBB, T. C. (1942): *J. Infect. Dis.* **71**, 228.

HARRIS, J. S. & KOHN, H. I. (1941*a*): *J. Pharmacol.* **73**, 383. — (1941*b*): *J. Biol. Chem.* **141**, 989.

HAWKING, F. (1937): *Lancet*, **2**, 1019.

HENRY, H. & STACEY, M. (1943): *Nature, Lond.*, **151**, 671.

HENRY, R. J. (1943): *Bact. Rev.* **7**, 175.

HIRSCH, J. (1941–2): *C.R. ann. Arch. Soc. Turque Sci. phys. nat.* **10**, 115. — (1942): *Studien über die mikrobiologischen Grundlagen der Sulfanilamid-Therapie*. Istanbul: Kenan Basimevi. — (1943): *Istanbul Seririyati*, **25**, 3.

HOFFMANN, C. E. & RAHN, O. (1944): *J. Bact.* **47**, 177.

JOHNSON, F. H., BROWN, D. & MARSLAND, D. (1942): *Science*, **95**, 200.

JOHNSON, F. H., EYRING, H. & KEARNS, W. (1943): *Arch. Biochem.* **3**, 1.

KNIGHT, B. C. J. G. (1936): *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 210.

KOHN, H. I. & HARRIS, J. S. (1941): *J. Pharmacol.* **73**, 343. — (1943): **77**, 1.

KUHN, R., MÖLLER, E. F., WENDT, G. & BEINERT, H. (1942): *Ber. dtsch. chem. Ges.* **75** (B), 711.

KUHN, R., WIELAND, T. & MÖLLER, E. F. (1941): *Ber. dtsch. chem. Ges.* **74** (B), 1605.

LANDY, M., LARKUM, N. W., OSWALD, E. J. & STREIGHTOFF, F. (1943): *Science*, **97**, 265.

LINEWEAVER, H. & BURK, D. (1934): *J. Amer. Chem. Soc.* **56**, 658.

LIPMANN, F. (1941): *J. Biol. Chem.* **139**, 977.

LOCKWOOD, J. S. (1941): *Chemotherapy*. Pennsylvania: Univ. Press.
LODGE, R. M. & HINSHELWOOD, C. N. (1939): *J. chem. Soc.* pp. 1683, 1692. — (1943): pp. 208, 213.
MCILWAIN, H. (1940*a*): *Brit. J. exp. Path.* **21**, 136. — (1940*b*): *Nature, Lond.*, **146**, 653. — (1941*a*): *Brit. J. Exp. Path.* **22**, 148. — (1941*b*): *Biochem. J.* **35**, 1311. — (1941*c*): *Nature, Lond.*, **148**, 628. — (1942*a*): *Lancet*, **1**, 412. — (1942*b*): *Brit. J. Exp. Path.* **23**, 95. — (1942*c*): **23**, 265. — (1942*d*): *Biochem. J.* **36**, 417. — (1942*e*): *Science*, **95**, 509. — (1943*a*): *Nature, Lond.*, **151**, 270. — (1943*b*): *Biochem. J.* **37**, 265. — (1943*c*): *J. Chem. Soc.* p. 322. — (1943*d*): *Brit. J. Exp. Path.* **24**, 203, 212. — (1943*e*): *Trans. Faraday Soc.* **39**, 359. — (1944*a*): *Nature, Lond.*, **153**, 300. — (1944*b*): *Biochem. J.* **38**, 97, 187. — (1944*c*): unpublished.
MCILWAIN, H., BARNETT, J. W. & ROBINSON, F. A. (1942): *Chem. Ind.* **61**, 96.
MCILWAIN, H. & HAWKING, F. (1943): *Lancet*, **1**, 449.
MCINTOSH, J. & SELBIE, F. (1943): *Brit. J. Exp. Path.* **24**, 246.
MCINTOSH, J. & WHITBY, L. E. H. (1939): *Lancet*, **1**, 431.
MANN, F. G. & WATSON, J. (1943): *J. Chem. Soc.* p. 606.
MARSHALL, E. K. jr., (1939): *Physiol. Rev.* **19**, 240. — (1941*a*): *Chemotherapy*. Pennsylvania: Univ. Press. — (1941*b*): *Ann. Rev. Physiol.* **3**, 643.
MARTIN, G. J. & FISHER, C. V. (1942): *J. Biol. Chem.* **144**, 289.
MATTI, J., NITTI, F., MOREL, M. & LWOFF, A. (1941): *Ann. Inst. Pasteur*, **67**, 240.
MELLON, R. R. & BAMBAS, L. L. (1937): *Proc. Soc. Exp. Biol., N.Y.*, **36**, 682.
MÖLLER, E. F. & BIRKOFER, L. (1942): *Ber. dtsch. chem. Ges.* **75**, B, 1118.
NETER, E. (1942*a*): *Amer. J. Surg.* N.S. **58**, 69. — (1942*b*): *J. Pharmacol.* **74**, 52.
PAGE, J. E. & ROBINSON, F. A. (1943): *J. Chem. Soc.* p. 133.
PAULING, L. & CAMBELL, D. H. (1942): *J. Exp. Med.* **76**, 211.
PETRO, J. (1943): *Lancet*, **1**, 35.
POOLE, E. A. & HINSHELWOOD, C. N. (1940): *J. Chem. Soc.* pp. 1565, 1573.
RAWLINGS, A. L., SWEET, L. A. & JOSLYN, D. A. (1943): *J. Amer. Pharm. Ass.* **32**, 11.
ROBERTS, E. C., CAIN, C. K., MUIR, R. D., REITHEL, F. J., GABY, W. L., VAN BRUGGEN, J. T., HOMAN, D. M., KATZMAN, P. A., JONES, L. R. & DOISY, E. A. (1943): *J. Biol. Chem.* **147**, 47.
ROBINSON, H. J. & GRAESSLE, O. E. (1942): *J. Pharmacol.* **76**, 316.
ROSE, H. & FOX, C. L. jr. (1942): *Science*, **95**, 412.
RUBBO, S. D., ALBERT, A. & MAXWELL, M. (1942): *Brit. J. Exp. Path.* **23**, 69.
SCHMELKES, F. C. & WYSS, O. (1942): *Proc. Soc. Exp. Biol., N.Y.*, **49**, 263.
SCHMELKES, F. C., WYSS, O., MARKS, H. C., LUDWIG, B. J. & STRANDSKOV, F. B. (1942): *Proc. Soc. Exp. Biol., N.Y.*, **50**, 145.
SCHULEMANN, W. (1932): *Proc. Roy. Soc. Med.* **25**, 897.
SELBIE, F. R. (1940): *Brit. J. Exp. Path.* **21**, 90.
SEVAG, M. G. & SHELBURNE, M. (1942): *J. Bact.* **43**, 421.
SHANNON, J. A. (1943): *Ann. N.Y. Acad. Sci.* **44**, 455.
SILVERMAN, M. & EVANS, E. A. jr. (1943): *J. Biol. Chem.* **150**, 265.
SNELL, E. E. (1941): *J. Biol. Chem.* **139**, 975; **141**, 121. — (1944): **152**, 475.
SPRAY, G. H. & LODGE, R. M. (1943): *Trans. Faraday Soc.* **39**, 424.
STACEY, M. & SCHLÜCHTERER, E. (1939): *Nature, Lond.*, **143**, 724.
STAMP, T. C. (1939): *Lancet*, **2**, 10.
STEARN, A. E. & STEARN, E. W. (1924): *J. Bact.* **9**, 491.
VALKO, E. I. & DUBOIS, A. S. (1944): *J. Bact.* **47**, 15.
VOEGTLIN, C. (1925): *Physiol. Rev.* **5**, 63.
WIGGERT, W. P. & WERKMAN, C. H. (1939): *Biochem. J.* **33**, 1061.
WILSON, G. S. (1922): *J. Bact.* **7**, 405.
WOOD, W. B. jr. (1942): *J. Exp. Med.* **75**, 369.
WOODS, D. D. (1940): *Brit. J. Exp. Path.* **21**, 74.
WOOLLEY, D. W. (1944): *J. biol. Chem.* **154**, 31.
WORK, T. S. (1940): *J. Chem. Soc.* p. 1315.
WYSS, O. (1941): *Proc. Soc. Exp. Biol., N.Y.*, **48**, 122.
WYSS, O., GRUBAUGH, K. K. & SCHMELKES, F. C. (1942): *Proc. Soc. Exp. Biol., N.Y.*, **49**, 618.
WYSS, O., RUBIN, M. & STRANDSKOV, F. B. (1943): *Proc. Soc. Exp. Biol., N.Y.*, **52**, 155.
WYSS, O., STRANDSKOV, F. B. & SCHMELKES, F. C. (1942): *Science*, **96**, 236.
YORKE, W. (1940): *Trans. Roy. Soc. Trop. Med. Hyg.* **33**, 463.
ZAHL, P. A., HUNTER, S. H. & COOPER, F. S. (1944): *Proc. Soc. Exp. Biol., N.Y.*, **55**, 4.

BACTERIAL GROWTH

By C. N. HINSHELWOOD, Physical Chemistry Laboratory, Oxford

(*Received* 22 *August* 1944)

The problem of the bacterial cell throws down a challenge to the physical chemist. On the one hand, bacteria in their small way show most of the mysterious characters of life. On the other hand, they are of a very simple structure, and although their internal composition is uniform neither physically nor chemically, the inhomogeneities are of a fine-grained and subtle kind (Knaysi & Mudd, 1943; Piekarski, 1939; Robinow, 1942). Although the surface layer of the cell possesses remarkable properties, of great importance to what goes on inside, it is best regarded as a boundary surface with characters determined by the laws of colloid chemistry.

The cell contains many enzymes, the substrates, intermediates and products of which diffuse about and may enter or leave it. The enzymes themselves are synthesized as the cell grows. All the cell material adds to itself by accretion of fresh units. It consists largely of proteins and other macromolecules, which are known to be well adapted to the building up of fibrous, laminar or three-dimensional network structures. We know that macromolecules can grow, often by the addition of free radicals; we know that the complex structures formed by them may be permeable to other substances; that a macromolecule consisting, for example, of protein united to polysaccharide could form part, with its two ends, of two different regions of crystalline order, united by an amorphous region of bent and tangled chains. We know that the surface of a macromolecular structure, or the internal surface of a network, could be the seat of heterogeneous catalytic reactions; and we might begin to picture the bacterial cell as an assemblage of more or less regular macromolecular structures, either separated by fluid, or possibly linked by chemically polyfunctional tie-molecules. Protein chains or networks will show a great variety of pattern, and the chemical properties will vary from region to region. The enzyme reactions which can occur in different parts of the cell may well be a function of the molecular pattern prevailing in the macromolecular network at the points in question. The organization of the cell may well depend, in part at least, upon the size and sequence of these various regions, and upon the nature of the intermediates which diffuse from one to the other. It is of the greatest interest to inquire which of the characters of the bacterial cell follow from the general proposition that it is the seat of a sequence of chemical reactions linked in space and time.

We know something of the laws of chemical kinetics, of surface reactions, chain reactions and the kind of reaction by which macromolecules are built up. Chemical kinetics is familiar with reactions linked in a time sequence; the addition of the spatial dimension and the related concentration gradients introduces a complexity into the problem which must lead to new phenomena. Whether these help in understanding biological behaviour is a problem of great interest.

It might be objected at the outset that even the bacterial cell is too complex to be studied as a whole. Much valuable work has been done on isolated enzyme systems. This mode of approach can, however, by itself hardly be enough. To understand the organization of a cell, the cell must be observed in its actual working. We wish to know not only the chemistry of the individual reactions, but the physicochemical mode of their linking; to study not merely isolated themes, but the laws by which the symphony is constructed.

In the following pages we shall consider what light is thrown on this problem by the study of three matters: first, the separation of the bacterial growth cycle into its various phases, and the factors which influence these; secondly, the linking of cell processes in relation to morphological changes, and to the influence of drugs on the organism; and thirdly, the adaptive changes undergone by cells in new media, or in the presence of antibacterial agents. All these will be considered primarily in connexion with bacterial growth rates.

I. THE BACTERIAL GROWTH CYCLE

(1) *General.* As is well known the growth of bacterial cultures can be divided into four main phases: the lag phase; the phase of logarithmic increase, during which the number of organisms is related to the initial number by the approximate law, $n = n_0 e^{kt}$; the stationary phase; and the death phase. In the logarithmic phase the number of cells doubles in equal increments of a period called the mean generation time (m.g.t.). The transition from the logarithmic to the stationary phase is usually rather sudden. Before considering some of the factors which control these phases, it will be convenient to outline a few general considerations about the kinds of reaction sequence which may be important in the cell.

We suppose that the essential basis of most

enzymes is a particular protein pattern. Further, it seems clear that a given enzyme may operate either on substrates supplied direct by the medium, or on intermediates which are themselves the products from an earlier enzyme in a temporal sequence. These intermediates may on occasion be very labile. Any stage in the series of processes may possibly be reversible, though this is not likely to be true of most of the stages. Any of the intermediates may be lost from the cell by diffusion, or otherwise wasted: there will be a competition between use by an enzyme and loss from the cell. The products from later members of the enzyme sequence may be fed back to contribute, together with substrates from the medium, to the working of earlier members.

In general we may regard any reaction involving cell substance and substrates as an enzyme reaction. Two major types must, however, be distinguished: *type A*, reactions not directly linked with growth; *type B*, reactions directly linked with growth and taking place according to the scheme:

enzyme pattern + substrate
= extended enzyme pattern
+ products, possibly used in other syntheses.

In type A the enzyme has a purely catalytic function. As has been pointed out (Davies & Hinshelwood, 1943; Hinshelwood & Lodge, 1944), the important kind of reaction must be of type B, since in fact all the enzymes reproduce themselves in the course of cell growth. The fundamental scheme given above has analogies in other parts of chemistry: for example, in the catalytic decomposition of arsenic hydride by arsenic, which yields more arsenic and sets free hydrogen. In general, when any ordered structure expands by addition of new units there is a decrease in free energy which can compensate the increases involved in the formation of other active products. The expansion of protein patterns in macromolecular assemblages may well occur according to this principle.

When the cells of a stable bacterial culture are in active growth the various enzymes must be expanding in a constant ratio. The rates of the various type B processes must therefore be equal and a steady state established for the whole sequence of reactions. If the ordered sequence is interrupted by the cutting off of the supply of a substrate, or by the inactivation of certain enzymes, the cells will enter the so-called stationary phase, during which various things may happen. Unstable intermediate products may decompose, diffusible ones will be lost, and the enzyme patterns themselves may suffer some degeneration, whether by chemical decomposition or by processes analogous to protein denaturation. If the cells are returned to a medium where the necessary substrates are supplied and toxic substances are absent, then the steady sequence of reactions characteristic of full growth will not necessarily be restored immediately. Enzymes may have to be rebuilt, and the diffusible intermediates must accumulate again so as to restore the concentration gradients necessary for the maintenance of the spatio-temporally co-ordinated reaction scheme. These things must happen during the lag phase.

(2) *The lag phase.* The work of earlier writers such as Penfold is described in textbooks. A summary of more recent investigations is given by Winslow & Walker (1939). Here we shall only refer to those more recent papers which bear upon the particular aspects we are going to consider.

During the lag phase we expect, first, the building up, by type A reactions, of the necessary threshold concentrations of some essential intermediates; secondly, the rebuilding of enzymes which have suffered structural degeneration; and thirdly, the re-establishment of the steady state, the second and third processes probably involving reactions of type B.

As regards the substances necessary for the onset of growth, we know that certain quite definite compounds are often required. Some of these may be very simple. One of the simplest is carbon dioxide. When an air stream freed from all carbon dioxide is passed through cultures in an artificial medium, growth may be delayed indefinitely, as shown by Gladstone, Fildes & Richardson (1935). The carbon dioxide affects not only the lag but the subsequent rate of growth, as shown for *Bact. lactis aerogenes* by Dagley & Hinshelwood (1938) and for pneumococci by Kempner & Schlayer (1942). The lag phase of *Bact. lactis aerogenes* is an artificial medium containing glucose, and a phosphate buffer is indefinitely lengthened as the concentration of magnesium ions present is reduced towards zero (Lodge & Hinshelwood, 1939). Magnesium, of course, must be provided from without, but the cells themselves can build up a critical concentration of carbon dioxide by fermentation reactions.

Usually, however, more complex substances play the important roles. Some bacteria are limited in their synthetic powers and will not grow unless they are provided with specific growth factors. For example, most of the pathogenic streptococci need ready-made glutamine (McIlwain, Fildes, Gladstone & Knight, 1939): in the absence of air *Staphylococcus aureus* in an artificial medium needed uracil, and, according to Richardson (1936), none of twenty-one similar compounds could replace it. This same organism also needs a potential source of the —SH group provided in an organic compound (Fildes & Richardson, 1937). Numerous other examples are known (Fildes & Richardson, 1935; Fildes, 1938; McIlwain, 1940; Gladstone, 1937). Sometimes bacteria, as ordinarily isolated, require definite organic compounds for growth, but can be trained, by gradual reduction of supply, to dispense with them. For example, *Bact. typhosum* normally requires tryptophan but can be trained to derive its nitrogen from ammonia (Fildes, Gladstone & Knight, 1933).

All bacterial protoplasm contains tryptophan units: some bacteria can be trained to build them up, others cannot (Fildes & Knight, 1933). Certain species can synthesize them if given indole (Fildes, 1940). The mechanism of the training phenomenon will be considered later.

In view of the variety of compounds which have appeared in the guise of growth factors—glutamine, uracil, nicotinic acid, thiamin and a number of amino acids—and especially in view of the graded series of requirements which may be run through on training, e.g. tryptophan, indole, ammonia, it looks as though the only difference between the exacting and the non-exacting types of bacteria is the precise point in the series of reactions at which operations begin. Sometimes the sequence of linked reactions begins with the utilization of such simple compounds as carbon dioxide and ammonia, and the products of the earlier syntheses diffuse from one region of the cell to the next, there to be utilized in the later stages of the sequence. In the so-called exacting species the necessary enzymes for the earlier stages are either absent—when no growth can occur without the external supply of relatively complex intermediates—or are inactive until they have been developed by the training process.

The intermediates may not be used or passed on as such: they may merely be the source of active fragments—possibly free radicals—readily participating in macromolecular chain growth reactions. This would explain in a general way observations such as that of Gladstone (1937), who found that an increase in the ability of *Staphylococcus aureus*, brought about by training, to synthesize alanine units was accompanied by the ability to dispense with valine, leucine and histidine, which originally had been demanded in the medium.

The quantitative study of the lag phase fully confirms the idea that bacteria which can utilize the simplest sources of substance build up their protein by way of intermediates which can sometimes diffuse from the cells into the medium, and which, therefore, can certainly pass from one internal department of the bacterium to be dealt with by another, in the way envisaged. In this connexion various phenomena are observed which are of considerable interest in themselves, and which will now be described.

An example which has recently been studied in some detail by Lodge & Hinshelwood (1943) is that of *Bact. lactis aerogenes*, which grows well in synthetic media and will utilize ammonium salts as the sole source of nitrogen. The relation between lag and inoculum age was studied. With glucose and asparagine as carbon and nitrogen sources the lag increased steadily with age over the first few days. With ammonium sulphate as nitrogen source, however, a phenomenon referred to as *early lag* appeared: very young inocula showed a considerable lag, which fell, as they aged, nearly to zero, and then rose again (*late lag*). The early lag was completely removed by sterilized filtrate from fully grown cultures, showing that the cells use a diffusible growth factor, which during the lag is built up to a threshold concentration. With older inocula enough is transferred in the medium to supply the needs, but with very young ones little is carried over, and it must be manufactured. The question arises whether perhaps all bacteria which use very simple compounds are dependent upon small stores of more complex substances handed on from one growth cycle to the next with successive inocula. In the present example this is not so, as can be shown by separating the cells of the inoculum from the parent medium before inoculating. The threshold concentration of the growth factor can be made up in two ways: by synthesis during the lag and by transfer from the parent medium. The greater the number of cells which constitute the inoculum, the sooner they can build up in the medium the necessary concentration for the use of all. This explains a marked dependence of the lag on the number transferred, an effect often noted, but not quantitatively studied. On the basis of simple assumptions a fairly satisfactory quantitative relation can be worked out giving the lag as a function of the actual number of cells on the one hand and of the volume of the original medium on the other hand, which are transferred with the inoculum. With amino acids as nitrogen sources, there is no early lag effect, and no influence of inoculum size.

There is an obvious relation between this study of the lag phase and observations which have been made in other connexions on diffusible co-enzymes. Examples are the growth activator for *Bact. coli* described by Sahyun, Beard, Schultz, Snow & Cross (1936), the diffusible co-enzymes postulated in connexion with deaminases in *Bact. coli* by Gale & Stephenson (1938) and in connexion with the lactic acid dehydrogenase of *Bact. coli* by Yudkin (1937). Woods & Trim (1942) find with the deaminases of *Clostridium Welchii* a variation of activity with dilution strongly suggesting a diffusible co-enzyme, which in the earlier stages of growth can be lost from the cells.

So long as there is no change in total cell substance, type A reactions, as defined above, are the only kind which can occur. They play the important part presumably in the early part of the lag phase. On the other hand, type B reactions may well occur in the final stages and herald the transition to the logarithmic phase. This transition is in fact not sudden. With most bacteria the size increases towards the end of the lag (Winslow & Walker, 1939), and there is an increase in metabolic activity greater than in proportion to the cell substance (Mooney & Winslow, 1935). Huntingdon & Winslow (1937) showed that the order of events was, first, increase in metabolic activity, then in cell volume and finally division.

As cells age, the lag increases and, under conditions

where the culture remains viable, may settle down to a more or less steady limit. At this stage it cannot be entirely removed by the addition of filtrate from fresh cultures.

The activities of individual enzymes, studied in washed suspensions of cells which are not multiplying, have been shown to depend upon the age at which these cells were collected for test from the growing culture. The activity rises to a maximum and then declines (Gale, 1940; Wooldridge & Glass, 1937; Woods & Trim, 1942).

Even where all these effects of ageing cannot be explained by loss of diffusible factors from the cells, they might be ascribed to the decay of labile intermediates. But it seems much more likely that the enzyme substance itself ages, and that reactivation occurs during the lag by the formation of new substance. One possible mechanism for the ageing is by the recombination, at the boundary of the enzyme structure, of the free valencies necessary for the propagation of polycondensation reactions. Neutralization of too many may involve death, but of a limited number merely inactivation.

The different enzymes in a cell may age at different rates: in which case it is conceivable that the original balance is never exactly restored on further growth, or that on repeated subculture the relative activities show secular changes, giving the appearance of cyclical variations in the bacteria.

One of the more important actions of antibacterial drugs has been shown to be specifically on the lag phase. Moreover, it has been shown that in the adaptive variation of bacteria to resist drugs or to utilize new food sources, it is the variation of the lag phase which may play the major part. These effects are considered more fully in a later section.

Rather meagre information is available about the relation between the lag and the concentration of the various constituents of the growth medium. Bacteriologists have usually been content to record qualitative observations of total growth after a given time. Thus it is doubtful whether the major factor has been mean generation time, lag, or total population which the medium can support. All these may vary independently according to circumstances. With *Bact. lactis aerogenes* deriving its carbon from glucose in a phosphate buffer the lag increases as the glucose concentration is increased: it is not much affected by the *p*H of the medium within limits. It is shorter with amino acids as nitrogen sources than with ammonium sulphate. More extensive data of various kinds are needed.

To summarize the processes which occur during the lag phase of a bacterial culture, there must be regeneration of inactivated enzymes and a fresh accumulation of labile or diffusible intermediates responsible for carrying the chain of reactions from one part of the cell to another. Reactions of type A occur first, without increase in cell substance, then reactions of type B, with linked working and expansion of the enzymes. Finally, when all the enzymes are active, and the necessary concentrations and concentration gradients of all the intermediates are established, the cell is ready for the steady period of co-ordinated activity which constitutes the logarithmic phase.

(3) *The logarithmic phase.* In this phase growth follows the approximate law $n = n_0 e^{kt}$, each cell dividing after a time which is on the average $0{\cdot}693/k$. Log n plotted against t gives a straight line.

The law is only an approximation, but is often a very good one. The reasons why, on the one hand, it lacks the character of an absolute law, and why, on the other, it is on the whole a very good approximation are themselves interesting. It lacks absolute character for the following principal reasons. First, the concentration of foodstuff in the medium changes as the cells grow. Secondly, products formed by the cells slow up growth as they accumulate. Thirdly, not every cell formed by the division of a parent survives to divide itself: thus, although the logarithmic law might be followed by the number of living cells, it will not be exactly followed by the total number. Fourthly, the conditions for cell division vary somewhat according to the state of the medium and therefore according to the stage of the growth cycle: thus the actual size of the cells changes as growth proceeds, and, although the logarithmic law might describe either total number or total bacterial mass, it could not do both simultaneously. The law works well in practice, however, because there is a considerable range over which these factors are unimportant. First, the rate of growth is usually almost independent of substrate concentration over fairly wide ranges. Secondly, there is usually a range of concentrations where toxic substances have little effect, but on the higher side of which their action increases very steeply. Thirdly, the mortality under favourable conditions may be very low. Fourthly, the rapid change in cell size tends to occur only at the extreme ends of the growth cycle. It will be convenient to give some further consideration to these various factors.

The relation between growth rate and concentration of medium constituents is in general represented by a curve which rises more or less linearly at low values of the concentration and then bends sharply round to become nearly independent of it. Only at very low concentrations of the normal food materials does the mean generation time show very serious increases (Penfold & Norris, 1912; Dagley & Hinshelwood, 1938; Lodge & Hinshelwood, 1939). This result can be easily understood if the relevant reactions are regarded as surface reactions with rates determined by the adsorbed substrate. If the adsorption isotherms follow the well-known Langmuir type, then the surfaces will be nearly saturated except when the concentration is low. The picture of a series of cell regions where reactions occur at rates determined by a Langmuir isotherm is found useful in interpreting other growth phenomena, e.g. in

connexion with adaptation: this will be referred to in a later section.

During growth the *p*H changes and toxic products accumulate. With *Bact. lactis aerogenes* it has been shown that the marked effect of *p*H is exerted mainly on the total population which the medium can support and that over a wide range the mean generation time is little influenced. The influence of an antiseptic is well known to vary steeply with its concentration, and is often expressed in terms of a power of the latter. Such a formula can hardly have a theoretical interpretation: but is empirically justified over limited ranges. The form of relation will be discussed later, but all that matters here is to note that in a quite narrow region of the growth curve the toxic action of products may rise from negligible proportions to complete inhibition. The result can be an almost abrupt transition from logarithmic to stationary phase.

According to Wilson (1922) with some organisms, even during the logarithmic phase, the percentage of new cells which are viable seldom exceeds 90% of the total. Kelly & Rahn (1932), however, made direct observations on the division of individual cells of various bacterial species and found that under favourable conditions all cells continued to divide. Even with Wilson's figure the departure from the logarithmic law for total numbers would not be serious.

The changes in cell size have been studied by Henrici and others (see review by Winslow & Walker, 1939), and more recently discussed by Hinshelwood & Lodge (1944), who have considered the influence on the mean generation time of morphological changes in an extreme case. Usually the changes are not such as to have much effect on the logarithmic phase.

The various factors which cause departures from the logarithmic law usually increase from negligibility to importance in a rather short period when the bacterial count is becoming greater than a certain limit: the food concentration falls, the toxic products accumulate, changes in the medium may make mortality more serious, and division conditions begin to be affected, all at an increasing rate as the numbers themselves increase in geometrical progression. There is then a transition to the stationary phase.

It is interesting to note that the bacterial mechanism can function satisfactorily over a wide range of speeds. *Bact. lactis aerogenes*, for example, will grow continously and well in media in which the mean generation time at 40°C. varies from 18 min. to over 200 min. There does not seem to be very much systematic information about growth rates in different media. Koser & Rettger (1919) examined the influence of different amino acids on micro-organisms and found no marked differences in availability. Gordon & M'Leod (1926) classified amino acids as indifferent, favourable or inhibitory. Sahyun *et al.* (1936) noted variations in the rate of utilization of various amino acids and wide differences in total growth. Often in the literature the effects on lag, growth rate and total population are not distinguished. Lodge & Hinshelwood (1943) measured the growth rates of *Bact. lactis aerogenes* in a series of amino acids with glucose as carbon source. The total population varied widely from one amino acid to another: sometimes depending also upon the aeration of the solution. The growth rates, on the other hand, all fell within a comparatively narrow range. The total population seemed to be in part controlled by the formation of an inhibitor removable by cellular oxidation. The earlier stages in the utilization of the amino acids seemed in general not to be rate-determining for the total sequence of growth reactions. The influence on growth rate of varying the carbohydrate was greater, the rate in glucose being about double that in lactose. Rahn & Richardson (1942) find that aeration of bacterial cultures affects total population but not necessarily growth rate.

(4) *The stationary phase.* Numerous views have been expressed about the reasons for the cessation of growth. Most of them are correct within the appropriate limits. On the whole it appears that the main factors are, first, exhaustion of substances necessary for growth, secondly, accumulation of toxic products and, thirdly, adverse *p*H. A quantitative study (Lodge & Hinshelwood, 1939), has shown that any one of these, according to circumstances, may in turn become the limiting factor. The influence of aeration on total population seems to be connected with the removal of toxic products by oxidation or otherwise. When medium exhaustion is the limiting factor, the total population becomes directly proportional to the initial concentration of a given food material in the solution.

One might have expected that growth inhibitors would affect the mean generation time alone, leaving the final total population unaltered. But the actual fact is otherwise. Inhibitors, such as alcohols, lengthen the mean generation time, i.e. slow down growth, but they also reduce the total final population in almost the same proportion. This can, however, be given a simple explanation in terms of the kinetics of the growth curve (Poole & Hinshelwood, 1940). This state of affairs stands in sharp contrast with what is found on variation of *p*H. With *Bact. lactis aerogenes* adverse *p*H reduces the final population almost to zero before any reduction occurs in the actual rate of the sparse residual growth. *p*H presumably affects growth by controlling the ionization of amphoteric protein structures. Again, one might expect *rate* of growth to be chiefly affected, which is not what is found. Therefore it seems likely that some intermediate substance plays a part. This is formed at a rate dependent upon *p*H and attains to some transient equilibrium concentration in the cell. The actual rate of growth is independent of this concentration over a wide range. At an adverse *p*H

the lowered production rate leads to a lowered reserve: thus when toxic substances accumulate in the medium they are able to slow up growth sooner, the stationary phase setting in rapidly when the supply of the intermediate substance falls.

(5) *The phase of decline.* If the medium is not renewed the cells gradually die. The rate of death is, of course, much increased by the presence of antiseptics. Only one aspect of the death of bacteria will be considered here, namely, the law giving the variation with time of the number of survivors in a declining population. This law raises a problem of fundamental interest, and has been the subject of much controversy. It is usually written in the form: *number of survivors* at time t = original number $\times e^{-lt}$, where l is a constant.

This form is similar to that giving the amount of unchanged substance in a unimolecular reaction. The similarity has unfortunately led to much misunderstanding. It will therefore be best to begin by stating the conditions under which such a law would hold, assuming it to be experimentally valid, which is not universally admitted. They are: first, that the death of a given cell is not influenced by the number of other survivors (as it would be, for example, if there were competition for residual food material), and secondly, that the chance of death in a short interval of time, dt, is independent of the previous history of the cell. No implications whatever about unimolecular reactions exist. The events conditioning death may be as complex as we please. The implications of the law are very general. For example, if a large number of individuals were to continue throwing dice until each had thrown six sixes in succession, then the number still in the game at any time would be given by the law in question. The law would still apply in essence, though it would be distorted in form, if some system of handicapping the individuals were introduced into the game. This case might be contrasted with that of a large number of freshly wound clocks, where the law giving the number still running at any time is not even roughly logarithmic. The number running would remain near the total for a time and then fall rapidly. The question is: to which of these two types of law does the decline of a living cell population more nearly correspond? One would have expected the time factor to be of prime importance, and the chances of death to increase with exposure to the adverse environment. If this were so, and the logarithmic law were still followed, it would be an accident, depending upon a particular initial inequality in the powers of individual resistance. Inequalities among members of the population are inevitable, but it is rather difficult to see how there could be an initial distribution of just the kind required, which would involve a maximum proportion not of average, but of the least resistant, individuals. At this point the problem is complicated by the fact that the law is not exactly obeyed: in the initial stages of decline there are often departures in the sense that would allow for the presence of some individuals more sensitive than the mean, but give the maximum in the intermediate region, which is statistically much more likely. (For discussion of data see Clark, *General Pharmacology*.) Nevertheless one has the impression that in some examples the logarithmic law is too good an approximation to be accounted for simply by the initial distribution, though it would follow at once if death were to some extent independent of previous history. The question must be regarded as open, but it is of great interest in connexion with cell organization. If death in presence of a poison, though conditioned by the presence of that poison, is independent of the time of action, then there would have to be some way in which the cell becomes sensitized to it at a given moment. Some accidental conjunction of events—as complex as we please—would be the signal for the final breakdown of the organization by which the cell had hitherto resisted the poison. Before speculation goes further in this matter it would be desirable to be more sure of the real status of the logarithmic law. There is no doubt about the importance of the issues it raises. A similar problem arises in connexion with the killing of bacteria by radiation: here the view has been advanced that quanta of radiation act as missiles, and that the chance event is the encounter between one of these and some sensitive region in the cell. But the statistical implications of the law are less specialized than this.

II. CELL DIVISION AND CELL MORPHOLOGY

Little is known about the nucleus of the bacterial cell: and there is no evidence that division is accompanied by mitotic phenomena resembling those which occur in some other kinds of cell. In this section we shall simply consider division from the point of view of the behaviour of the whole cell.

Kelly & Rahn (1932) observed the division of large numbers of individual cells of different species of bacteria, and found quite large variations in the fission times. The following figures are typical of their observations on *Bact. lactis aerogenes* at 30° C. The numbers of cells with fission times in the various ranges were:

Min. ...	5–10	10–15	15–20	20–25	25–30	30–35
	0	1	11	25	42	97
Min. ...	35–40	40–45	45–50	50–55	55–60	
	65	45	20	8	2	

The daughter cells from an early division were not consistently different from the average of all the cells: thus what was being witnessed was not a selection of faster growing cells from a mixed strain, but a random fluctuation.

The random fluctuation is more likely to be due to internal than to external causes, and one might

suggest a variable delay in the reorganization of the cell contents just before division, the process being in some measure analogous to the variable nucleus formation in a supercooled liquid. The analogy must not be pressed too far: the essential thing is that both the cell reorganization and the nucleus formation in the liquid are thought of as dependent upon some conjunction of multiple probabilities.

Ordinarily the division probability becomes very high as soon as the cell attains to a more or less standard length. In the above example very few cells have a fission time more than twice the mean value. The result is that the distribution of lengths among cells is usually not a very wide one, though the mean length may vary systematically over the growth cycle (Henrici, 1922, 1923). There may, however, be circumstances where division is delayed, while elongation of the cell continues normally, so that long filamentous or snake-like organisms are formed, fifty or a hundred times the ordinary length. These long cells are sometimes observed to occur spontaneously, and sometimes to be produced in the presence of foreign substances such as dyes or drugs (Ainley Walker & Murray, 1904; Tunnicliff, 1939; Gardner, 1940).

A fairly detailed study has been made of the conditions of long cell formation with *Bact. lactis aerogenes* (Hinshelwood & Lodge, 1944). In normal cultures the size distribution at any moment does not show more scattering than could be accounted for by the different ages of the cells present, and it appears that the rate of elongation is the main factor determining the moment of division. When certain strains are transferred from one medium to another long cells appear during the first few subcultures, before adaptation to the new medium is complete. The size distribution is now quite different. In some circumstances the number of cells with a length greater than a given multiple of the average follows a law of the same form as that expressing the number of molecules in a gas which traverse a distance more than a given multiple of the mean free path. The probability of division is now so much lower that the elongation rate is no longer the limiting factor. The cells go on growing until some internal conjunction of events allows division to occur. The formation of an exceptionally long cell depends upon the probability that this conjunction of events is delayed to an unusual extent: this is statistically analogous to the combination of events which permits a molecule to survive without collision an abnormally long free path. Various observations of the influence on size distribution of inoculum age, inoculum size, the presence of filtrate from older cultures, serial subculture, and on the specific influence of drugs can be correlated with the aid of the following hypothesis. There are two separate factors, one of which is diffusible into the medium, responsible respectively for the elongation and for the division of the cells. The latter factor is consumed or diluted in the process of division and its formation may be specifically accelerated or retarded by the presence of foreign substances. When cells are transferred to a new medium, the rates of formation of these two factors, which were originally balanced, may become out of balance. If the elongation factor is formed more readily than the division factor, then the filamentous forms appear. On successive subculture the balance is restored by a mechanism which will be discussed under the heading of adaptation.

III. INFLUENCE OF DRUGS ON BACTERIAL GROWTH

In this section we shall only be concerned with drug action as an instrument for exploring the co-ordination of reactions in the cell economy. (For more general aspects see Faraday Society Discussion: 'Modes of Drug Action,' 1943.) The first important result to note is that various drugs may have separate and specific actions on lag, mean generation time, division probability, and death-rate. The distinction between bacteriostatic and bactericidal agents is emphasized in modern work: though the difference in the end is one of degree rather than one of kind.

The conventional law of antiseptic action sets the influence of the drug proportional to a power of its concentration, which may be quite a high one. This, however, is in need of rational interpretation. A third or fourth power law could hardly correspond to any physico-chemical theory. The action of drugs on the lag or the mean generation time can be represented by formulae of rational form which also express rapid variations with concentration. For example, the action of alcohols on the growth rate of *Bact. lactis aerogenes* is given by $k=k_0-ap$, where k_0 is the growth rate constant in the absence of alcohol, and k that found at concentration p, a being a constant. The variation of a with the chain length of the alcohol suggests that each CH_2 group attaches itself independently to some element of the internal structure of the cell (Tilley & Schaffer, 1926; Stiles, 1936; Meyer, 1937; Dagley & Hinshelwood, 1938). Thus the linear form depends simply upon the shape of an adsorption isotherm governing the equilibrium between the toxic substance in solution and that taken up by the centres which it inactivates. Sometimes there is a threshold concentration below which the drug has little effect. This is because some of the drug may be taken up and neutralized by the organism, a process in which the internal concentrations of various intermediates of metabolism may sometimes play their part. When there is an initial tolerance of this kind the growth rate may vary with concentration according to an expression of the form

$$k=k_0-a\,(p-p_0).$$

From this it follows that

$$d\log k/dp=-(k_0/k-1)/(1-p_0/p).$$

In the region where p is only slightly greater than p_0 this coefficient may have a very large numerical value, so that if k is expressed empirically in the form $k = 1/p^x$, then x would come out with a high value. This would seem to be the essential basis of the power laws sometimes used to express antiseptic action.

An interesting special case may occur where apparently the cells grow at a normal rate up to a critical concentration of the drug above which they refuse to grow at all. This can be explained as follows: the drug acts primarily on the lag but has little effect on mean generation time. At higher concentrations the drug lengthens the lag to such an extent that the cells all die before growth begins: once growth does begin, as it can at lower concentrations, it is normal. Examples of different types of rate-concentration relation have been discussed by Poole & Hinshelwood (1940), in the light of the view that the drugs interfere with the cell processes at varying stages.

The differential action of drugs on different phases of bacterial growth has been established in various ways. There is a close connexion between this matter and the successive deactivations of resting *Bact. coli* described long ago by Quastel & Wooldridge (1927). But the most direct evidence comes from the studies made in recent years by Fildes, Woods, McIlwain, Gladstone and others. The underlying idea is that specific inhibitory actions may be exerted by substances which are structurally related to the normal metabolites of the cell. These substances are taken up by enzymes in competition with their normal substrates, but, being useless for further parts of the reaction sequence, block the whole process of growth. Fildes (1940) showed that the antibacterial action of mercury may be specifically and quantitatively neutralized by compounds containing —SH groups, and supposes that the toxicity of the mercury depends upon the blocking of such groups in cell metabolites or enzymes. Woods (1940) suggested that sulphonamides interfere with the use of aminobenzoic acid in the cell. McIlwain (1940) studied inhibition by pyridine-3-sulphonic acid and its amide, which are supposed to act in competition with the nicotinic acid of normal metabolism: the inhibitors are taken up by and put out of action the centres which normally would use the nicotinic acid. The amino-sulphonic acid analogues of some natural amino acids have an analogous effect (McIlwain, 1940). Other examples are analogues of pantothenic acid (McIlwain, 1942) and indoleacrylic acid, which is thought by Fildes (1941) to act as an inhibitor at some stage in the synthesis of tryptophan. A somewhat similar idea emerges from the work of Gladstone (1939) who finds that excess of one amino acid may put out of action the centres needed for the synthesis of another related one, and thus stop growth unless that other one is independently provided.

In the light of these general ideas we may now consider some quantitative work on the antibacterial actions of diaminoacridine (proflavine) and of methylene blue on *Bact. lactis aerogenes*. One of the principal effects of both is greatly to prolong the lag period (Davies, Hinshelwood & Pryce, 1944). As explained above, the lag of young cultures is removed by the addition of filtrate from grown cultures, and, since the action of the drugs is strongly antagonized by the addition of extra filtrate, a natural hypothesis is that they prevent the synthesis or utilization of this particular growth intermediate. The relation between lag and drug concentration is interesting. With proflavine there is an initial tolerance and then a very rapid increase of lag with concentration. With methylene blue the initial tolerance is followed by a much slower increase. The relations can be more or less quantitatively accounted for by quite simple assumptions. (*a*) The drug interferes with the production by one enzyme of a sequence of an intermediate utilized by a later member of the sequence. (*b*) The rate of working of the second enzyme is related to the concentration of the intermediate by an equation of the form of a Langmuir isotherm. (*c*) The rate of increase of lag with drug concentration depends upon the completeness with which the drug can remove the intermediate: with proflavine there is something like a quantitative titration, but with methylene blue there is an equilibrium, hence the rapid rise with one and the slower rise with the other once the tolerance limit is passed. (*d*) The tolerance itself arises from the form of the Langmuir isotherm: if the working range of intermediate concentration corresponds to the flat part of this isotherm, some degree of removal is possible before the rate of working of the second enzyme begins to be affected. This idea is helpful in connexion with the adaptation of the cells to resist the drugs (see next section).

While dealing with the relation between bacterial growth and drug concentration, reference should be made to the special form of curve met with, for example, in the action of sulphonamides on *Bact. lactis aerogenes*. The growth rate falls steadily to a limit below which no amount of the drug will reduce it. There seem to be some growth centres which are immune or inaccessible (Davies & Hinshelwood, 1943). The phenomenon of what are virtually alternative modes of growth, of which this is probably an example, seems to be of some general importance: it is discussed by Lodge & Hinshelwood (1944) but cannot be considered further here.

Another important specific action of certain drugs is on the cell division probability. Where this is much reduced without a corresponding lowering of elongation rate, long filamentous cells appear. Numerous references to long cells appear in the literature (as mentioned above); with *Bact. lactis aerogenes* our information is beginning to be a little more systematic. Phenol, sulphonamide and methylene blue never seem to provoke their formation, but *m*-cresol, tertiary butyl alcohol and proflavine readily do. The

formation is always a function of the concentration of the drug and of the age of the inoculum, just as with the spontaneous formation discussed earlier in this paper, and the results can in general be interpreted by postulating separate elongation and division factors, *L* and *D*, independently modifiable by the drug (Lodge & Hinshelwood, 1943; Spray & Lodge, 1943; Hinshelwood & Lodge, 1944; Davies *et al.* 1944). Spray & Lodge found that cells which had given filaments in meta-cresol would sometimes show abnormally slow growth for many subsequent generations in their normal medium, some difficultly reversible change in cell balance having occurred.

Some of the most remarkable phenomena which occur with bacteria in the presence of drugs are those of adaptation: these will be considered in the following section.

IV. BACTERIAL ADAPTATION

The subject of bacterial variation is a vast one, of which only those aspects will be discussed here which can be brought into relation with growth phenomena. The characteristics of bacteria change easily. The classification of divisions within a given species is sometimes a matter of difficulty, and a continuous series of intermediate forms may exist (e.g. Fildes, 1927). The important species characters remain unchanged (Virtanen, 1934). Coliform organisms, for example, never change into streptococci, nor do organisms occur which can be regarded as intermediate forms between the two species. Yet the subdivisions of the coliform species itself merge imperceptibly into one another, and it is at least possible that they change into one another. (For studies of the coliform group see e.g. Parr, 1939; Nyberg, Bonsdorff & Kauppi, 1937; or Sievers, 1937.) What is certain is that on repeated subculture in new media, on growth in presence of drugs, or even by what in ignorance we call spontaneous variation, bacteria of a given strain acquire new characters, or at least undergo changes in the quantitative balance of old ones. These changes are often reversible, but it must be emphasized that their amplitude is limited. Some characters are never lost or gained, and the range of quantitative change is seldom indefinitely extended, though occasionally a minor character may be more or less permanently lost (e.g. Neri, 1940).

When variations or adaptive changes occur we must be dealing with one of two things: either a shift in the numerical balance of populations which from the start were mixtures of substrains, the variation being simply the selection of the substrain best adapted to flourish under given conditions; or else the actual modification, under the influence of these conditions, of the character of the individual cells.

Widely different views have been held on this matter (see e.g. Mellon, 1942; Hadley, 1937; Todd, 1930). The balance of evidence seems to be definitely in favour of the view that actual modification of individual cells occurs, though no doubt selection can be and will be superimposed on this when modified and unmodified cells exist together. That the essential role in bacterial adaptation or variation is played by cell modification is supported by the following arguments: (1) Variant strains arise from cultures which sprang originally from a single cell (e.g. Torrey & Montu, 1936; Dombrowsky, 1936); thus, even if selection were the major factor, the original type had at some stage to become heterogeneous, an assumption which itself admits the modification of the individual cell. (2) Adaptation is often a multiple process, different characters developing at different rates, as found by Reed (1937) with *Serratia marcescens* and by Lodge & Hinshelwood (1943*b*) and Davies & Hinshelwood (1943) with *Bact. lactis aerogenes*. According to the rigid selection hypothesis, not two, but a whole spectrum of non-interconvertible substrains would have to be postulated here. (3) When *Bact. lactis aerogenes* is partially trained to resist sulphonamide, the immunity is lost if the cells are grown several times in the normal medium. According to the selection idea we can easily understand the survival of resistant strains in the presence of the drug, but why the non-resistant strains should regain their ascendancy in its absence would be unexplained. (4) When *Bact. lactis aerogenes* is trained to resist the action of diaminoacridine, the degree of adaptation is exactly graded to correspond to the concentration of drug in which the organism was grown. This quantitatively graded response can be much more naturally explained by a shift in the enzyme balance of all the cells, than by selection between inherently immune and inherently sensitive strains. (5) Even stable variants will usually revert to the original forms under appropriate treatment. It is true that some variants or trained organisms are very stable indeed. For example, when *Bact. lactis aerogenes* has been trained by growth in presence of proflavine, twenty passages through the normal growth medium in the absence of drug do not weaken the immunity. Indeed, Hadley (1937) suggests, for the special example of rough colony forms of bacteria, that reversion only occurs when the variant has not been thoroughly isolated. Ease of reversion may indeed become less the more thoroughly the training of an adapted form has been carried out: for example, after a few passages through sulphonamide *Bact. lactis aerogenes* easily loses its immunity; after many, the immunity will outlast any number of passages through normal media. This might seem like an argument for the view that one strain is selected, and a competing strain eliminated with more or less efficiency in the two cases. The argument will not hold, however, because, under appropriate conditions, the stabilized new form, from which by hypothesis any competitor has been eliminated, will in fact often revert. The stable rough colony forms referred to above sometimes

revert when passed through an animal; and the sulphonamide-fast strains of *Bact. lactis aerogenes* lose their immunity when grown in presence of proflavine.

All the various changes can be more simply assumed to depend, in their initiation at least, upon a modification either of the enzyme balance of the cell or of the mode of cell division, caused by the medium, by the drug, or by other factors.

Before discussing this in further detail it will be useful to say something about what is called bacterial dissociation. This is a particular manifestation of variation affecting the form of the colonies in which the organisms grow on solid media. It is perhaps a little unfortunate that so great a proportion of the literature on variation deals with this phenomenon, because we do not know at all well on what other properties of the cells the colony form (smooth, rough, mucoid) depends (though it seems that rough colonies are often associated with longer cells—see Hadley, 1937; Roelcke & Intlekofer, 1938). However, a brief summary of the principal results must be given since these foreshadow in many respects those which are found in more quantitative studies of growth phenomena.

A given culture may yield colonies of a single form or of a mixture of forms; on subculture these colonies may breed true or revert. The production of variants is influenced by medium, chemical reagents, age of culture, X-rays and other agencies. It occurs with most if not all species of bacteria: sometimes the variant forms revert easily, sometimes they are stable (e.g. Faragó, 1934). The reversion is sometimes sudden, erratic and apparently spontaneous. Even stable variants will often revert on special treatment. The different colony forms are linked in varying degrees with other properties such as morphology, pathogenicity and biochemical reactions (see e.g. Scholtens, 1937).

Some observers have thought the different variants to be different parts of a bacterial life cycle. But potentially reversible processes succeeding one another in an imperfectly controllable manner are bound to produce such an effect, and on the whole the life cycle is not generally believed in (Rettger & Gillespie, 1933, 1935).

The influence of the environment in determining variation is stressed by some writers, for example, Flynn & Rettger (1934) and Gillespie & Rettger (1939), while others stress the inherent power of the cell to give variant forms on division. Bunting (1940, 1942), working with the pigment-forming *Serratia marcescens*, which gives red, pink and white colonies, found that when the strain was kept multiplying logarithmically and samples were plated out from time to time, a definite equilibrium ratio was reached with about 97% dark reds and 3% pale pinks. The rates of variation were considered to express the 'probability of the occurrence of specific intracellular events'; the environmental influence seemed to be small. Lewis (1934), with *mutabile* strains of *Bact. coli*, concluded that all colonies contained some proportion of variants, and Deskowitz (1937) found that with *Salmonella aertryke* under fixed conditions of environment the ratio of variant to parent type was constant, the conclusion being that each cell, by an inherent property of its protoplasm, could produce a variant form once in a given number of divisions. Parr & Robbins (1942), working on variants of coliforms, find that some fail at first to utilize citrate as a carbon source, but always contain among their progeny some few cells which can. They speak of 'mosaics of potentiality'. The general impression left by many papers on bacterial dissociation is that of multiple possibilities which present themselves at the moment of cell division. A rather different perspective is offered by studies on the adaptation of bacteria to grow in new media or to immunize themselves to the action of drugs. Here one receives the impression that a very important part is played by actual changes in the enzyme balance occurring in the cell, not so much at the moment of division, as during growth. The two aspects are not at all irreconcilable and, indeed, a shifting enzyme balance on the one hand, and a multiple division mode, on the other, may be equally important and even interrelated factors.

Quantitative studies on the growth rate throw some light on these matters. In adapting themselves to change of medium, bacteria will sometimes only utilize new sources of carbon or nitrogen at maximum rate after training by repeated growth in the new environment. During the adaptation morphological changes may manifest themselves. Serial subculture in presence of drugs may give progressively increasing growth rates. The limits of adaptation are definite. On the whole there is no good evidence that training to resist higher temperatures can occur. With *Bact. lactis aerogenes* adaptation will occur to many drugs, but none at all to phenol. Subculture in presence or absence of glycerol causes the rate of growth in the latter to rise and fall, but only between restricted limits.

The nature of the adaptation occurring when *Bact. lactis aerogenes* is trained to use a new carbon or nitrogen source has been investigated by Lodge & Hinshelwood (1944). In the new medium the curve of bacterial number against time is initially composite. At a given point there is a transition from a slower to a more rapid mode of growth. In the first subcultures in the new medium the transition occurs late in the growth, but as training by successive subculture proceeds, the transition occurs earlier and earlier. What appears to happen is that growth in the new medium occurs by two mechanisms, a less efficient one with a short lag, and a more efficient one with a longer lag. Training consists in the progressive shortening of the lag phase of the growth mechanism giving the greater rate in the new medium. This adaptation only occurs during actual cell multi-

plication in the new medium, not by mere sojourn there. The phenomena can be interpreted as follows. Various possible reaction sequences can lead to the final products necessary for the building of the cell substance. In a new medium a new reaction sequence will correspond to the optimum rate. This new sequence will, however, involve different intermediates, and hence will have a long lag, until enzymes have been built up in the cell, capable of yielding the necessary new intermediates. While this change in enzyme balance is occurring growth still goes on partly by the old reaction sequence, which, however, in the new medium functions below its optimum rate. Analogous phenomena occur during the training of *Bact. lactis aerogenes* to resist sulphonamides.

The great significance of the lag phase in adaptation phenomena comes clearly to light in experiments on the training of *Bact. lactis aerogenes* to resist the inhibitory action of proflavine, which we shall consider here in some detail rather than attempt to summarize the literature on drug resistance in general (for earlier literature see Borman, 1932).

The initial tolerance of *Bact. lactis aerogenes* to small amounts of proflavine, the subsequent rapid rise of lag with drug concentration, the neutralization of the drug action by filtrate from grown cultures and the possibility of adaptation have already been referred to. After several cultures in presence of proflavine the lag is reduced almost to that which would be found in the absence of the drug. Under optimum conditions the immunization can be complete after a few cell divisions, though no progress is made towards it during the lag phase itself. Thus it seems that the enzyme balance only changes in the sense favourable to drug resistance during the actual period of increase of cell substance. The rapidity with which drug resistance can be established was observed by Fleming & Allison (1927) in the action of lysozyme (from tears and other secretions) on the organism *M. lysodeiticus*. The firmness with which the acquired resistance is retained was also noticed in this connexion, and is found with *Bact. lactis aerogenes* and proflavine; though here quantitative study reveals a more complex picture. The curves of lag against proflavine concentration for a series of cultures trained at successively higher concentrations constitute a family, all similar in shape, but with increasing initial tolerances, the increase in the latter being over a considerable range directly proportional to the concentration at which the training occurred. (The theory of this will be referred to later.) If the training concentration is below a certain limit, the acquired immunity may persist after many passages in the normal medium. When, however, the cells have been trained to resist specially high concentrations, then the immunity decreases spontaneously, on simple passage through the normal medium, until it falls to a limit, much below the maximum, but much above that of the untrained cells. Theoretical considerations, based upon the idea of adaptation by modification of the enzyme balance in the cell, led to the prediction that immunity might be completely destroyed if the trained culture were grown in presence of other antibacterial agents than the one to which it had been trained. This is confirmed: proflavine-trained cells lose their immunity on subculture in presence of cresol or phenol, sulphonamide-trained cells, on subculture in presence of proflavine (Davies *et al.* 1944).

In all these examples, one of the important modes of adaptation is the progressive shortening of the lag phase on successive subculturings. This raises an incidental point: the lag itself is sometimes spoken of as a period of adaptation, which is unfortunate. The lag is a period during which necessary intermediates are built up, and in this it resembles the induction period of certain chemical reactions. True adaptation is the modification, during growth, of the mechanism by which such intermediates are built up, or of other cell mechanisms.

Changed medium constituents in general, or drugs specifically, change the relative balance of the factors determining cell elongation and division respectively. Filamentous forms, as already explained, may appear under the new conditions if the balance is changed in a particular way. On successive subculture, however, the division factor becomes adapted and the filaments disappear. With *Bact. lactis aerogenes* filaments formed on change of medium, or under the influence of proflavine, disappear after a few subcultures (but those provoked by meta-cresol persist as long as the drug is present, and certain other conditions are fulfilled). This sort of effect is partly responsible for the idea that life cycles exist. From what has been said, it seems that random observations of variations in morphology should not be given this interpretation.

The idea that drugs may intervene at specific points of reaction sequences is supported by the observations on adaptation. *Bact. lactis aerogenes* when trained to proflavine is immune also to methylene blue and vice versa: but when trained to these drugs it is not immune to sulphonamide. This shows that the two former attack predominantly at one point of a sequence, while the latter attacks at another. Phenol, on the other hand, provokes no adaptation at all.

The theoretical aspect of adaptation phenomena is one of the most interesting problems which bacteriology presents to physics and chemistry. The types of theory which can be made fall under three heads: (1) those based upon selection, (2) those based upon the idea of a changed enzyme balance, (3) those based upon the idea of a change in the mode of cell division or in some heredity-bearing mechanism in the cell.

Reasons have already been given for excluding theories wholly based upon selection. Types (2) and (3) need not be mutually exclusive, since changes

in enzyme balance can affect modes of division and vice versa.

We shall first consider the question of changing enzyme balance. Enzymes in bacteria have for some time been classed as constitutive or adaptive, the latter being those which are only called into existence in presence of the appropriate substrate (Dubos, 1940). In the light of what has been said above, it seems that the distinction is largely a question of relative values of lag periods, and is a function of age and other conditions. Yudkin (1938) envisages reversible systems of enzymes and precursors of enzymes such that, if the enzyme is removed by combination with an inhibitor or otherwise, the equilibrium between it and the precursor is shifted to help to restore the balance.

A considerable number of facts can be correlated if we bear in mind that, according to the fundamental equation of type B reactions (p. 151) enzymes have two functions: to reproduce their own substance in cell growth, and to provide substrates for other enzymes linked in sequence with them. A simple model can be imagined of two linked enzymes, the product from the first constituting the substrate of the second, which it reaches by diffusion, with the alternative possibility of loss into the medium. Equations expressing the rates of growth of the two enzymes can be written down and from them various conclusions can be drawn (Davies *et al.* 1944), namely: (*a*) If a drug acts by reducing the availability of the intermediate substrate for the second enzyme, adaptation should occur. (*b*) The adaptation only occurs during actual growth, but under optimum conditions may be complete after a few divisions. (*c*) Spontaneous loss of adaptation may occur with very different degrees of readiness. (*d*) Induced reversal of the adaptation should be caused by specific drugs.

One very important factor in this simple theory is the relation between the rate of working of an enzyme and the concentration of its substrate. That assumed is similar to a Langmuir adsorption isotherm. If the working range corresponds to the flat part of this isotherm, then the concentration of the substrate can be reduced to some extent without effect on the rate of working. This may lead to a range of tolerance to the drug. If the working range corresponds to the rising portion of the curve, the working of the enzyme will be retarded by a reduced supply of substrate: adaptation can then occur by the expansion of the supplying enzyme to make good the deficit. This expansion may be automatic, because cell division may await the building up of a standard amount of the second enzyme, and by the time this is achieved, the first will have increased beyond its normal amount. If trained cells are returned to normal conditions, the intermediate substrate may be supplied far in excess of need by the expanded enzyme. In some cases this can cause an automatic reversal of the adaptation, but if the excess merely shifts the working range farther along the flat part of the isotherm, there may be a neutral equilibrium with no readjustment of balance. Agents which actively change the relative rates of growth of the two enzymes may, however, cause loss of training.

Ideas of this sort can go a considerable way in explaining the facts about adaptation: but it is unlikely that they tell the whole story. Changes in the enzymatic activities of drug-trained cells have actually been observed. Sulphathiazole-resistant strains of *Staphylococcus aureus* are stated to show increased aminobenzoic acid production (Landy, Larkum, Oswald & Streightoff, 1943) and *Pneumococcus* variants obtained under sulphonamide influence are stated to show gradations in peroxide formation, fermentation reactions and virulence (McKinney & Mellon, 1941).

According as the enzyme most directly influenced by the drug occurs in the whole reaction sequence before or after the one responsible for a particular function tested, the influence of drug adaptation on that function would be expected to be different. This accounts for the fact that sometimes groups of properties vary in a dependent and sometimes in an independent manner.

The view that variation and adaptation depend upon changes in some sort of nuclear organization must be given its due weight. A schematic picture of the sort of thing which might be involved is suggested by Reed (1933) who envisages an occasional abnormal mode of splitting of certain cell constituents. Ordinarily the nucleus splits into two equal parts, but on occasion an asymmetry develops and one of the cells possesses a new kind of nucleus: this is the variant cell. The time is not ripe to discuss the physical chemistry of cell division. The advantage of a view like this, however, is that it would account easily for the occasional great stability of variants, and would explain a certain incalculability in their occurrence.

The hypothesis of alternative modes of division is likely enough in itself, and will probably be worked out further. It seems likely that division mode and enzyme balance both play their parts, and indeed one probably interacts with the other. As we have seen already, the hypothesis of elongation and division factors invoked to explain morphological changes shows how a changed enzyme balance reacts upon the moment of cell division; and the moment of division might well react upon the mode of division. On the other hand, if there are alternative division possibilities, resulting differences in nuclear constitution would certainly react upon the enzyme potentialities of the cells.

We have seen, therefore, that a great many phenomena can be related, in principle, to changes in the quantitative distribution, during growth, of the existing forms of bacterial substance. Enzymes in linked sequences may wax and wane in amount according as the supply of their intermediates varies;

division may be advanced or retarded; the modes of cleavage of nuclear material may vary. But all these things happen without any change in the fundamental protein patterns of the basic cell constituents. To change these would require an upheaval of a much more profound kind. This is doubtless why variation occurs between well-defined limits only: why biochemical characters may change quantitatively and great morphological differences may be produced, and yet the species, as far as ordinary observation of its essential characters goes, remains inviolate. The reason why the fundamental protein patterns do not change would seem to be this: the free energy which allows the growth of highly active chemical units is provided in part at least by that which is liberated when new fragments build themselves on to old ones after the manner of the growth of crystals. The running down of potential energy is greatest when the new units form an orderly array with the existing ones. The generation of a new species from an old one would demand processes quite unlike this, and which, without the source of free energy, would be of almost transcendent thermodynamic improbability.

V. SUMMARY

(1) The structure of the bacterial cell is simple: the various regions of enzyme activity may be macromolecular networks of specific pattern which grow by polycondensation reactions. Intermediate substrates diffuse from one enzyme to another in the cell, so that the activity of the bacterium is at least in part a sequence of chemical reactions linked both temporally and spatially in definite relations. Various growth phenomena can usefully be looked at from this point of view. A fundamental scheme of enzyme synthesis is considered. (2) The various phases of the bacterial growth cycle are considered. During the lag, among other events, concentrations of diffusible intermediates are built up: these intermediates may be lost from the cell (which accounts for various observations, including an effect of inoculum size on lag). The starting-point of the reaction series varies with the species of bacterium. When all enzymes are active, and all the intermediates are present with stationary concentration gradients, the logarithmic phase begins. The conditions determining the validity of the logarithmic growth law, the factors governing the setting in of the stationary phase, and the law expressing the decline of the viable population are discussed in the light of experimental data. (3) Cell elongation and cell division are factors separately modifiable by media or by drugs, with resulting changes in morphology. In particular the formation of cells of abnormal length is considered. 'Life cycles' are not accepted. (4) The laws expressing the influence of drug concentration on bacterial growth are discussed. Drugs may attack specifically different points in the reaction sequence: they may influence lag and growth rate independently. The action of proflavine on a coliform organism is considered in some detail. (5) The adaptation of bacteria to use new sources or to resist drugs is thought to depend upon modifications induced in individual cells during growth, and not to be explicable simply by a selection of substrains. (6) Various groups of experimental facts are described and the hypothesis put forward that certain adaptation phenomena arise from a modification of the relative proportions of different enzymes, occurring during the expansion of the cell substance. The possibility that variants arise from multiple modes of cell division is also considered: change of enzyme balance and multiple modes of division are probably interrelated effects. (7) The spontaneous and induced reversion of drug adaptation is considered. (8) The limits of adaptation and the reason for the stability of the main species characters is discussed.

VI. REFERENCES

AINLEY WALKER, E. W. & MURRAY, W. (1904): *Brit. Med. J.* **2**, 16.
BORMAN, E. K. (1932): *J. Bact.* **23**, 315.
BUNTING, M. I. (1940): *J. Bact.* **40**, 57, 69. — (1942): **43**, 593.
DAGLEY, S. & HINSHELWOOD, C. N. (1938): *J. Chem. Soc.* pp. 1930, 1936, 1942.
DAVIES, D. S. & HINSHELWOOD, C. N. (1943): *Trans. Faraday Soc.* **39**, 431.
DAVIES, D. S., HINSHELWOOD, C. N. & PRYCE, J. M. (1944): *Trans. Faraday Soc.* **40**, 397.
DAVIES, D. S., HINSHELWOOD, C. N. & PRYCE, J. M. (unpublished work).
DESKOWITZ, M. W. (1937): *J. Bact.* **33**, 349.
DOMBROWSKY, K. H. (1936): *Z. Bakt.* **137**, 160.
DUBOS, R. (1940): *Bact. Rev.* **4**, 1.
FARAGÓ, F. (1934). *Z. Bakt.* **133**, 139.
FILDES, P. (1927): *Brit. J. Exp. Path.* **8**, 219. — (1938): **19**, 239. — (1940): **21**, 67, 315. — (1941): **22**, 293.
FILDES, P., GLADSTONE, G. P. & KNIGHT, B. C. J. G. (1933): *Brit. J. Exp. Path.* **14**, 189.
FILDES, P. & KNIGHT, B. C. J. G. (1933): *Brit. J. Exp. Path.* **14**, 343.
FILDES, P. & RICHARDSON, G. M. (1935): *Brit. J. Exp. Path.* **16**, 326. — (1937): **18**, 292.
FLEMING, A. & ALLISON, V. D. (1927): *Brit. J. Exp. Path.* **8**, 214.
FLYNN, C. S. & RETTGER, L. F. (1934): *J. Bact.* **28**, 1.
GALE, E. F. (1940): *Biochem. J.* **34**, 392, 846.
GALE, E. F. & STEPHENSON, M. (1938): *Biochem. J.* **32**, 392.
GARDNER, A. D. (1940): *Nature, Lond.*, **146**, 837.
GILLESPIE, H. B. & RETTGER, L. F. (1939): *J. Bact.* **38**, 41.
GLADSTONE, G. P. (1937): *Brit. J. Exp. Path.* **18**, 322. — (1939): **20**, 189.
GLADSTONE, G. P., FILDES, P. & RICHARDSON, G. M. (1935): *Brit. J. Exp. Path.* **16**, 335.
GORDON, J. & M'LEOD, J. W. (1926): *J. Path. Bact.* **29**, 13.
HADLEY, P. (1937): *J. Infect. Dis.* **60**, 129.
HENRICI, A. T. (1922): *Proc. Soc. Exp. Biol., N.Y.*, **20**, 179. — (1923): **21**, 215, 343, 345.
HINSHELWOOD, C. N. & LODGE, R. M. (1944): *Proc. Roy. Soc.* B, **132**, 47.
HUNTINGDON, E. & WINSLOW, C.-E. A. (1937): *J. Bact.* **33**, 123.
KELLY, G. D. & RAHN, O. (1932): *J. Bact.* **23**, 147.
KEMPNER, W. & SCHLAYER, C. (1942): *J. Bact.* **43**, 387.
KNAYSI, G. & MUDD, S. (1943): *J. Bact.* **45**, 349.

KOSER, S. A. & RETTGER, L. F. (1919): *J. Infect. Dis.* **24**, 301.
LANDY, M., LARKUM, N. W., OSWALD, E. J. & STREIGHTOFF, F. (1943): *Science*, **97**, 295.
LEWIS, I. M. (1934): *J. Bact.* **28**, 619.
LODGE, R. M. & HINSHELWOOD, C. N. (1939): *J. Chem. Soc.* pp. 1683, 1692. — (1943*a*): pp. 208, 213. — (1943*b*): *Trans. Faraday Soc.* **39**, 420. — (1944): (in the press).
McILWAIN, H. (1940): *Brit. J. Exp. Path.* **21**, 25, 136, 148. — (1942): *Biochem. J.* **36**, 417.
McILWAIN, H., FILDES, P., GLADSTONE, G. P. & KNIGHT, B. C. J. G. (1939): *Biochem. J.* **33**, 223.
McKINNEY, R. A. & MELLON, R. R. (1941): *J. Infect. Dis.* **68**, 233.
MELLON, R. R. (1942): *J. Bact.* **44**, 1.
MEYER, K. (1937): *Trans. Faraday Soc.* **33**, 1067.
'Modes of Drug Action' (1943): *Trans. Faraday Soc.* **39**, 319.
MOONEY, G. & WINSLOW, C.-E. A. (1935): *J. Bact.* **30**, 427.
NERI, F. (1940): *Z. Bakt.* **146**, 166.
NYBERG, C., BONSDORFF & KAUPPI, K. (1937): *Z. Bakt.* **139**, 13.
PARR, L. W. (1939): *Bact. Rev.* **3**, 1.
PARR, L. W. & ROBBINS, M. L. (1942): *J. Bact.* **43**, 66.
PENFOLD, W. J. & NORRIS, D. (1912): *J. Hyg., Camb.*, **12**, 527.
PIEKARSKI, G. (1939): *Z. Bakt.* **144**, 140.
POOLE, E. A. & HINSHELWOOD, C. N. (1940): *J. Chem. Soc.* p. 1565.
QUASTEL, J. H. & WOOLDRIDGE, W. R. (1927): *Biochem. J.* **21**, 148.
RAHN, O. & RICHARDSON, G. L. (1942): *J. Bact.* **44**, 321.
REED, G. R. (1933): *J. Bact.* **25**, 545. — (1937): **34**, 255.
RETTGER, L. F. & GILLESPIE, H. B. (1933): *J. Bact.* **26**, 289. — (1935): **30**, 213.
RICHARDSON, G. M. (1936): *Biochem. J.* **30**, 2184.
ROBINOW, C. F. (1942): *Proc. Roy. Soc.* B, **130**, 299.
ROELCKE, K. & INTLEKOFER, H. (1938): *Z. Bakt.* **142**, 42.
SAHYUN, M., BEARD, P., SCHULTZ, E. W., SNOW, J. & CROSS, E. (1936): *J. Infect. Dis.* **58**, 28.
SCHOLTENS, R. T. (1937): *Z. Bakt.* **139**, 467.
SIEVERS, O. (1937): *Z. Bakt.* **139**, 27.
SPRAY, G. H. & LODGE, R. M. (1943): *Trans. Faraday Soc.* **39**, 425.
STILES, W. S. (1936). *Introduction to Plant Physiology*, p. 81.
TILLEY, F. W. & SCHAFFER, J. M. (1926): *J. Bact.* **12**, 303.
TODD, E. W. (1930): *Brit. J. Exp. Path.* **11**, 368.
TORREY, J. C. & MONTU, E. (1936): *J. Bact.* **32**, 329.
TUNNICLIFF, R. (1939): *J. Infect. Dis.* **64**, 59.
VIRTANEN, A. I. (1934): *J. Bact.* **28**, 447.
WILSON, G. S. (1922): *J. Bact.* **7**, 405.
WINSLOW, C.-E. A. & WALKER (1939): *Bact. Rec.* **3**, 147.
WOODS, D. D. (1940): *Brit. J. Exp. Path.* **21**, 74.
WOODS, D. D. & TRIM, A. R. (1942): *Biochem. J.* **36**, 501.
WOOLDRIDGE, W. R. & GLASS, V. (1937): *Biochem. J.* **31**, 526.
YUDKIN, J. (1937): *Biochem. J.* **31**, 865. — (1938): *Biol. Rev.* **13**, 93.

CAMBRIDGE: PRINTED BY W. LEWIS, M.A., AT THE UNIVERSITY PRESS